

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

January 8, 2001, 09:20:44 ; Search time 2611.99 Seconds
(Without alignments)
2315.928 Million cell updates/sec

Title:
US-09-553-431-3

Perfect score:
1182

Sequence:
1 aagctgatactgcactcc.....tcaaaaaaaaaaaaaa 1182

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:

1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters:

2236266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1:
1: gb_ba1:
2: gb_ba2:
3: gb_cm:
4: gb_ov:
5: gb_ph:
6: gb_pl1:
7: gb_pl2:
8: gb_pl3:
9: gb_pr2:
10: gb_pr3:
11: gb_ro:
12: gb_sy:
13: gb_un:
14: em_fun:
15: em_hum1:
16: em_hum2:
17: em_in:
18: em_cm:
19: em_or:
20: em_ov:
21: em_pat:
22: em_ph:
23: em_pl:
24: em_ro:
25: em_sts:
26: em_sy:
27: em_un:
28: em_vl:
29: gb_ba3:
30: gb_in1:
31: gb_in2:
32: gb_in3:
33: gb_pl3:
34: gb_pr4:
35: em_ba1:
36: em_ba2:
37: em_hlg1:
38: em_hlg2:
39: em_hlg3:
40: em_hlg4:
41: em_hlg5:
42: em_hlg6:
43: em_hlg7:

44: em_hlg8:
45: em_hlg9:
46: em_hlg10:
47: em_hum3:
48: em_hum4:
49: em_hum5:
50: em_hum6:
51: gb_pr5:
52: gb_pr6:
53: gb_pr7:
54: gb_hlg1:
55: gb_hlg2:
56: gb_hlg3:
57: gb_hlg4:
58: gb_hlg5:
59: gb_hlg6:
60: gb_hlg7:
61: gb_hlg8:
62: gb_hlg9:
63: gb_hlg10:
64: gb_hlg11:
65: gb_hlg12:
66: gb_hlg13:
67: gb_hlg14:
68: gb_hlg15:
69: gb_hlg16:
70: gb_hlg17:
71: gb_hlg18:
72: gb_hlg19:
73: gb_hlg20:
74: gb_hlg21:
75: gb_hlg22:
76: gb_hlg23:
77: gb_sts1:
78: gb_sts2:
79: gb_vl1:
80: gb_vl2:
81: gb_pat1:
82: gb_pat2:
83: em_hlg0:
84: gb_hlg24:
85: gb_pr8:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1164	98.5	1164	7	AP251019
2	455.8	38.6	65958	6	AB009056
3	454	38.4	981	6	AB030278
4	205.8	17.4	118360	6	AF16114
5	186	15.7	150613	6	AB001684
6	185.4	15.7	194509	6	AB001129
7	163.8	13.9	200799	6	AF137379
8	163.8	13.9	200799	6	AF137379
9	159.6	13.5	22085	33	PM2245645
10	157.8	13.4	121524	6	AF041468
11	120	10.2	135638	29	SYCSLNG
12	95.4	8.1	14103	1	AE000712
13	94.8	8.0	673	6	AF149810
14	84.8	7.2	294250	2	AP001517
15	77.2	6.5	16486	1	AE001931
16	73	6.2	1900	2	BSMIND
17	73	6.2	4389	2	BACREMIN
18	73	6.2	5641	2	BACDIVERG
19	73	6.2	218410	2	BS0B0015
20	65.8	5.6	9966	1	AE002374
21	63.8	5.4	293181	2	AP001119
					AP001119 Buchnera

ALIGNMENTS

PLN 06-SEP-2000

Eukaryota; Viridiplantae; Charophyta
 Magnoliophyta: eudicotyledons; core eudicots; Asteridae; eastersterids
 II. Asterales: Asteraceae; Asteroideae; Heliantheae; Tagetes.

Analysis of carotenoid biosynthetic gene expression during petal development

Unpublished
2 (bases 1 to 1164)
Moehs, C.P., Tian, L., and DellaPenna, D.

TITLE Direct Submission
Submitted (03-APR-2000) Biochemistry, Univ. of Nevada, Reno, NV
JOURNAL 00557 USA

FEATURES	Location/Qualifiers
source	1. . 1164

```

2./organism="Tragetes erecta"
3./db.xref="taxon:13708"
4.
50. .937
6./codon_start=1
7./product="mind"
8./protein_id="AAG10431.1"
9./db_xref="GI:9971822"

```

BASE COUNT
ORIGIN

Query Match	98.5%	Score 1164	DB 7	Length 1164
Best Local Similarity	100.0%	Pred No. 7	3e-256	
Matches 1164; Conservative	0	Mismatches	0	Indels 0; Gaps 0

Score 1164; DB 7; Length 1164;

[illegible]

Oy 1081 gagaaatctgaatctgagcaatttcattgcattcaatcgaatcgtatcttcttgc 1140
 Db 1081 GAGAAATTTAGAAATTTGACCACTTTTCATTGATTCATTCATTAATTTGATTTCTTGGTTC 1140
 Oy 1141 agcagcaattcttaccacaaatc 1164
 Db 1141 AGTGAATGATTTTACTCAAAATC 1164

RESULT 2

LOCUS AB009056 65958 bp DNA PLN 09-AUG-2000
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MZF18.
 ACCESSION AB009056
 VERSION AB009056.2 GI:9758219
 KEYWORDS

ORGANISM

Arabidopsis thaliana (strain:Columbia) DNA, clone_1lb:Mitsui P1
 clone:MZF18.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsids.
 1 (sites)

REFERENCE

AUTHORS

Sato,S., Kaneko,T., Kotani,H., Nakamura,Y., Asamizu,E., Miyajima,N.
 and Tabata,S.

TITLE

Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 Sequence features of the regions of 1,456,315 bp covered by
 nineteen physically assigned P1 and TAC clones

JOURNAL

DNA Res. 5 (1), 41-54 (1998)

MEDLINE

98290546

REFERENCE

2 (bases 1 to 65958)

AUTHORS

Nakamura,Y

TITLE

Direct Submission

COMMENT

Submitted (27-NOV-1997) to the DDBJ/EMBL/GenBank databases.
 Yasukazu Nakamura, Kazusa DNA Research Institute, Department of
 Plant Gene Research, 1532-3, Yana, Kasarazu, Chiba 292-0812, Japan
 (E-mail: ynakamura@kazusa.or.jp, Tel:81-438-52-3935,
 Fax:81-438-52-3934)
 On Aug 9, 2000 this sequence version replaced gi:2656032.
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MZF18
 'Product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Grail-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University, and
 http://gremlin.zool.jastat.eu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-se
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-se/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is M001 and the 3' clone is M0E8.

FEATURES

Location/Qualifiers

CDS

1..65958
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="MZF18"
 /clone_1lb="Mitsui P1"
 join(355..1064,1116..5684)
 /note="gene_id:MZF18.17"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 /product="non-LTR retroelement reverse transcriptase-like

CDS

protein"
 join(8240..8326,8416..8695,8789..9161,9246..9406,
 9571..9698,9776..10024,10109..10399)
 /note="gene_id:MZF18.16"
 /codon_start=1

CDS

/evidence=not_experimental
 /product="(+) -delta-cadine synthase (d-cadine synthase)"
 /protein_id="BAB08719.1"
 /db_xref="GI:9758220"
 /translation="MGSEVNRPLADPPANIWDPLTFSKSDLGTEPEKSKTLKEA
 VKEAPMSSKANPNIENIKFTDALCRGKGVSYHEKDIYEDLQKSEFCDLDPPOKRGCD
 LITVGIITQVRFQFGLSADVFEKFGKNGKGLHVTDAVGMSTVEAMQMGVMA
 DIIDRLAFSRSHLEISRSRSPHAIIRKRNKPKYKGSIRITROTISTYEES
 CDPTLLEAFKIDFNLIQILHREELACVIRWNNHMEPKSVYTRHRTVEAYLMSGTY
 GCGPLKSLQKTANGYMOEAKWKLDYATFDEKVALISGVALIAMTFRMTD
 VAKIDAFEMLSHPRIRYASRISFTDIDISSYEFERHEVAVAGIDCYMOFGYSKE
 RAVEVMGNIYSDAKMDLQDELMPHVEFPLMRYLNSRVIVDFYRQADVTNPKLL
 KEHIVSLIETPTI"
 11597..12883
 /note="gene_id:MZF18.15"
 /codon_start=1
 /evidence=not_experimental
 /product="acetyl-CoA:benzylalcohol acetyltransferase-like
 protein"
 /protein_id="BAB08720.1"
 /db_xref="GI:9758221"
 /translation="MKLELSKEVIRKASPNHLOTLSLFDPLPSTVYSALFFIDD
 HSNQEDIORLKSLSQTLSEFLPLAGIQKGVKHNDEGALTEPARAELEFDFLRN
 PSDADLIQKIVSPDHADPEPTMLHVKVYFEKDFGFAVAVSNIKICDASLSTVC
 NNRASKGAYADVDPEFVGADYFPADIDISTEPEPLIVHETKSKTRFVSGSLMEIKL
 HKAIGNFFPLPKRESSEKMEFEYSKLOKTKOELNLRINDSDAKSVEAKRGI
 ASAMLSLCEISPEMETYVSSMCKNSFTYEAANGWCKPVVAVDSVDTQVYLMDSKD
 SEGVETARVLTPTDMAKYIHDSELLVATPSPSII"
 complement(1019113438..13580,13695..14419,14516..14753,
 14956..15283,15400..15634,15951..16167,16240..16311,
 16400..16533)
 /note="gene_id:MZF18.14"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08721.1"
 /db_xref="GI:9758222"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLN
 MNPGLVGTNLVLTVPMMFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVLPFGWILWIMISTNLFKSKWTRLSKYLK
 TTYEGPOGMNIVLITVPMKFLAVLSGVLIHRLNLSOSYLYECREMKLRITMGAVTAV
 GYVNICWAELEFPVTRASTILPLVGLSSESTYHIMLGHVSNFCELVITVPLIY
 AMINKMETFAMNPTVYVNLAGTAMVIGIAMVYSLPSFRKKELEFETVHNLGYL
 IPEVIVHGVSWPCMTLPNIFLEFIDRYLRLOSITKRSRLVASRLISDNLLEFST
 PGILVPTSLIFLHVPSTISKIOMHPETITSSNLEKDTLSRFLVROGSMOKLTFST
 SSIDSLSEVTEBGRPNLSPVSRHNSILIVSGSGITPFIISVIRELISOSONSTKL
 DVLLYCFKRIHDAFDLIFPLDMSADISRLNLRLEAVITBEDKPEPTTDHRLSL
 TKMFKRPLDPSIPVYGPNNFLMGVILSVYVMEFLITGITRYITIPVDINTGSI
 INFSTKGLMWMFGLSACIFSSSVFLMRKKKEDDEKRNQVSEETPTSGS
 WFGHRELESVYOSIVQATSVHFGSKPLKLTILBESSEVGVAVGCPRRMREY
 AKTSSGLAKNLHFEAISFNM"
 complement(1019120836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861)
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence=not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:975

Oy 209 ctgcgagagaacactccacgaagctcgcaatcagcgcgcgacgclggccccccagcacct
||||| | | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32248 CTCGCTCGTTACGGTTCCTCAGATTGTCCGCATTAGCGCGACTTGGTCTCCGTAACTC 32307

QY	269	gabcctctctcgtgtctcgaanaacccggtcaattacaccgttggttgaaattctcaagc	328
Db	32308	GATCTCTCTCAGGGTTAGAGAAATCGAGTCAATTAACATCTTCCGACGGTTATMAACGA	32367
QY	329	gattgcagaactcgaccgaagccctagtttcgttgataaaccgtgtgcgaattctcgattgctt	388
Db	32368	GATTGTCTCTCGATACAAAGCTCTGTGTACGTGATAGAGTTGGTCCGAATTTCCGAAATGCTTA	32427
QY	389	tgtatttcaaaacctagtgtaaaattgcttccttaggattttgggaaaaagcttagttgtg	448
Db	32428	TGTATATCTMAACCTAGATCCGAATCTTCGATGGGATTTGGTGTAAGACATTGGAAATGG	32487
QY	449	ctt----gattgcattaaagaatagagaagaagaattgcccggatttlaactatagattgt	505
Db	32488	CTTTGTGGATGCGTTTAAAACTAGAACCGGAAGTTCCACGCGATTTCAATCATCATGATTTGT	32547
QY	506	cctgcagatattgatgcgcgggttctaaaccgcgcattaaaccgcgttaagaccgtatla	565
Db	32548	CCGTGCGAATCGATCGCCGGATTTCAATACCGCATATTCTCCGGGATGAAACAGTTCTCG	32607
QY	566	gttaaacaccctggaattacatcgcatttagagagaatgagaataagtatacaggtctgtgaa	625
Db	32608	GTTACACACTCCGGATTTAAACACCGTTAAGAGATGCTGATAGGGTTTACCGGTTTGTAAGA	32667
QY	626	tgtgatagaatttagagatttaaaatgatattggaacagagttagaacgatttgtgaag	685
Db	32668	TCCGATCGAATTCAGACATATMAAGATGATTTGGAACAGGTGACAGACTGATTTGATTTAA	32727
QY	686	ggtgagaatatgtgttcaggtcttcttgatttcaagaagatgtttggatttgtcatctgttggt	745
Db	32728	GGAGAGGATATGATGTCTGATGTTGATGATGTCAGAGAAATGTGGGATTTGCATTTGGCTGTGT	32787
QY	746	g----atacccgagatctggaagtgtctgagtaacgaatagaaggtgttcctcgtgtgtg	802
Db	32788	GTAATTCCTCGAAGATTTCTGAGTTATTCGAAGCACGATCGAGGGTTTCCGCTGTTCTCG	32847
QY	803	aacaaagcttcggaacttagacagaatttgcatattgaagaagctgtccttggaagattggttg	862
Db	32848	AATAAGCCTCTACGCTTTCGGGGATTTGGCGCTTTGACACGCGCGCTTGGAACTCGTTTGAG	32907
QY	863	caagaatacatgaagcgtgtgtatgtgtgagaagaacacttaaaaaagaagaggaattcttcctg	922
Db	32908	CAAGTAGATGMAAGCGTTATGTGTGAGGAAGAACCTTAAGAAACGTGGCTTCTCTCT	32967
QY	923	ttttttgaggttagtgatcgaattcgttgatcg	957
Db	32968	TTCTTTTGGCGCTAATATGCAAAACGATCGCAACCG	33002

RESULT	3
AB030278	
LOCUS	981 bp mRNA PLN 11-PEB-2000
DEFINITION	Arabidopsis thaliana mRNA for MIND, complete cds.
ACCESSION	AB030278
VERSION	AB030278.1 GI:6759276
KEYWORDS	MIND.
SOURCE	Arabidopsis thaliana cDNA to mRNA.
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi. Arabidopsis. 1 (sites) Klim,M., Fujiwara,M., Kanamaru,K., Tanaka,K. and Takahashi,H. Arabidopsis thaliana mind homolog involved in plastid division Unpublished (1999) 2 (bases 1 to 981) Kim,M., Fujiwara,M., Kanamaru,K., Tanaka,K. and Takahashi,H. Direct Submission Submitted (22-JUL-1999) to the DDBJ/EMBL/Genbank databases. Nucleo Takahashi, Institute of Molecular and Cellular Biosciences, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-0032, Japan


```

gene
CDS
/AAAGTEKGIERTPEVLFMKPLD"
/complement(3568..4629)
/gene="psbA"
/complement(3568..4629)
/gene="psbA"
/codon_start=1
/product="D1 reaction center protein of photosystem II"
/protein_id="AAFA3791.1"
/db_xref="GI:7259490"
/translation="MTATLERRESANLWGRCFEITSTENRLYLIGFVIMIPCLLTA
ISVYIIAFAVAPVDIDGIREPYSGILAGNNITSGSVTSMNATGLHFYPIWEASL
DEWLYNGGPLYAVVHPLGLIACVMSREMLSPRLGPMIAVAYASAPAAATAVFLI
YPIGGSFSDGMLGISTENEMIVQAEHNIIMHPHMGVAGVGGSLFESAMHSLA
VTSLLIRETENSANAGYKFGQEEETENIVAAHFGPLIROYASFNRSLSHFEFLA
VMPVGIWTFAMGISTMAFNMGFNENQSVDSQGRVINTWADIIRANLGMVMEHR
NAHNEPDIASVEAPAVNG"
/complement(4754..6550)
/gene="ftsI"
/complement(4754..6550)
/gene="ftsI"
/codon_start=1
/product="putative peptidoglycan synthetase"
/protein_id="AAFA3792.1"
/db_xref="GI:7259491"
/translation="MNKKIKFGESRIYLVWILFFSGSLLGLFLYQVWKGSMULTK
AKNQITILNTPORRTICDRNGIPLAIDTAYDIFAHPLHFKSTEDIANELSTILN
LDVEYLONLFINNTGICIAHAPROANOIANKIEGIELVQHPKRYPYKQICADY
IGVNTLHEGQAGLESCQESLOLOSPEVVAIDGRGLALIDTAPSYDPNYYPFIE
FTDILOKASVLAIDYDGIKCKNAKGTIYIIDPYGALIALVTAPSYDPNYYPFIE
RKNMPVIDLYEPGSTFKPINMAIALEKAKIKNDPFTYQIOGLGNNVLETGELS
OPACKNSHLINTDVLNSNSVGNVHIIQRLAPETIYQIOGLGNNVLETGELS
SYSSLKILEFTSYNIESAVTSFGQGLAMTPKRLQALVACLSNGCNTIRPYLYQELF
IONERLFTNNIPONISLKKLTKVFSPTSEIYIDMLEVIFNGTSSCCLPQ
YRLGKTGTSOKHAROGYSTKHLITLFAIPINNPVYVILVIDEISPLSFSNT
ACPVVRSIIESLIRIKKIPSPITLTHNYCK"
/complement(6788..7477)
/gene="clpP"
/complement(6788..7477)
/gene="clpP"
/codon_start=1
/product="proteolytic subunit 2 of clp protease"
/protein_id="AAFA3793.1"
/db_xref="GI:7259492"
/translation="MPGIPKVAARIPEAVSYVDVYNNRYERILFLGEDLDEVA
NOLIGVAVFLNSEDVDTGIFTYINSFGSGNAGIGVDMIOHNVNVTTCMGLASM
ASFILAGCTGQGLMPFHARVMDHOPMGCGAKYAVVESVYKRLRELIALHYKR
TGSLEKIRKDMNDNMPROAKKEYGLIDHMTVNVNDELDELKQSDMLKLGKVNLT
NIERSNKSSEK"
/7825..9351
/gene="psbB"
/7825..9351
/gene="psbB"
/codon_start=1
/product="CP47 chlorophyll a apoprotein of photosystem II"
/protein_id="AAFA3794.1"
/db_xref="GI:7259493"
/translation="MGLPMYRVHTVVLNDPQRLISVHIMTGLVSGNAGSMAFVFLAV
FPSPDPVLPMMRQGMFVLPENTRLGISKSWGCMINDSITDPLMSTBEVAAHTHT
LAGMLFLASMHVWYWDLELFRDPRTGKPADLPIFGIHLFLSGLLCFGGAEHVTG
LFGPGIWSDPYGITGRQPIRPSWAGDFDPFNGGIAHSHTIAGIIGIAGLFLHS
VRSERFLYKALRMGNVEYVLSSTIAVMAAFVSGTMMYGAATPIELFGSTRQMD
LGFFNKEINKRVAQASIASGSIASEMSRIPKELATPYDITGNPARKGFLRAAMNGD
GIAAGNLGAHVAEKDKGRELFAVRMTPEFETFPVLLDKDGIADIPRRRESYSI
FOGVSAVAVAGGELDGVTEKDEPTVYKVARAOLGIEFEDNARKLSQCVPRSSRGV
FTGHLCPALIFPFGHIWMACTIFRDVPAIGIDPDIDQEVFGAFQKLDASTKQAV

```

```

gene
CDS
/db_xref="GI:7259494"
/translation="MEALVYTFPLVGLTGLIIFFAIFRPPRIIK"
/complement(9669..9800)
/gene="psbN"
/complement(9669..9800)
/gene="psbN"
/codon_start=1
/product="N protein of photosystem II"
/protein_id="AAFA3796.1"
/db_xref="GI:7259495"
/translation="MTATLITSTFSCLLIGLITGYSLYTSFGNASSELRLDPFEHED"
/9891..10112
/gene="psbH"
/9891..10112
/gene="psbH"
/codon_start=1
/product="10 kDa phosphoprotein of photosystem II"
/protein_id="AAFA3797.1"
/db_xref="GI:7259496"
/translation="MADTSQGRVYVGNFLKPLNSEYKVAPEGMTVLKGFMALEFA
VFLVITLIQYNASVLDGIPANWSSLKY"
/10445..11092
/gene="petB"
/10445..11092
/gene="petB"
/codon_start=1
/product="apocytochrome b6 of cytochrome b6/f complex"
/protein_id="AAFA3798.1"
/db_xref="GI:7259497"
/translation="MSKVYDMFDRLEIGIADITSKYVPPVNVNFCIGITPFCF
IMVAGSEAFMEFYRPTVTEAFASVOYIMDVNFGMLISIRKMSAMVLMILHVF
RYLTGCFKKPRELTVETGLVAVCTVSEVGTYSLEPMQVGVAVKIVTGPDAIV
IGAVLVELLAGVGVGOSTLRRYSLHFTVLPILTVFMIAHFLMKIKGIGSLP
/11202..11684
/gene="petD"
/11202..11684
/gene="petD"
/codon_start=1
/product="subunit IV of cytochrome b6/f complex"
/protein_id="AAFA3799.1"
/db_xref="GI:7259498"
/translation="MSIOKKPDLKDPVLRALAKGMSHNYGEPAMNDLLYTFPVC
IGTIGCLAVLAEPTMGEPRNPAPPELEIPMEFPVPOIIRVPIPKLGLVNLAA
GVPAIGLIVPFIESTINKFQNFRRVMAVTVLLIGVAVVWVGIGATLPIIDRSLHGF
/complement(11779..12192)
/gene="rps9"
/complement(11779..12192)
/gene="rps9"
/codon_start=1
/product="ribosomal protein S9"
/protein_id="AAFA3800.1"
/db_xref="GI:7259499"
/translation="MSVOKMPPILNYLGTGRKSAVARYVPGNGEVLINGLPCTNY
LOPNGSYLSVRSPLLETGLIEDNYDIIVAVAGGLTGQAEAIRLGAVALRLCIDISNR
HPLKKGFLTRDSRVKRRKYGGLKRAKAPQESKR"
/complement(12245..13195)
/gene="rpoA"
/complement(12245..13195)
/gene="rpoA"
/codon_start=1
/product="alpha subunit of RNA polymerase"
/protein_id="AAFA3801.1"
/db_xref="GI:7259500"
/translation="MSGNDLFPSTIYCTIESKIESPRNLVGRFLIEPLAIGGITVGN
LRLILIGESAITSYKIRGANNEISLFGIRSVLEILNLKELIVFRKSLDVQK
YLSIGPCVAVANMLQPLTSIEVDGGQYIATLSGNMLDMEFVINYGKGYOMADYAI

```

Query Match 17.4%; Score 205.8; DB 6; Length 118360;
 Best Local Similarity 60.3%; Pred. No. 9.3e-37;
 Matches 363; Conservative 0; Mismatches 237; Indels 15; Gaps 2;
 QY 229 agctgcgaatgcagccgacgttggtctaagtaaccctcgtctctcgtcgcga 288

```

Db 94002 ACTCGCTTTAAATGATGCTGATGCTGTTAAAGAAATTTAGATTTATTTAGACTTGA 93943
Oy 289 aaaccgylcaattacacacgctgctgtaagttctcaacgagctgcagactgcagcaagc 348
Db 93942 AAATGCTGAATTTATATACGCCATGAGACTTTTGAAGAGAAATGCTCTTGATCTAACG 93863
Oy 349 cctgctgctgtaagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 408
Db 93862 ATTATTTAGGCTAATACACATGAGATTTTACCTTTGCTGGCCATTTTAAACCAAGCA 93823
Oy 409 aaactgcctttagagcttgggagaaagccttagtctgctgctgctgctgctgctgctgct 468
Db 93822 AAGATATCATATTTAAACAAAGAAATATGGAATGCTT-----GTAGACTCTAT 93775
Oy 469 gcaagaaggttcccgatatttactatagatttctcgcagctgctgctgctgctgctgctgct 528
Db 93774 AAGCTTAAGAAATTTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 93715
Oy 529 cataccgcatctacacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 588
Db 93714 TGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 93655
Oy 589 atcgaagagctgcatagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 648
Db 93654 GATTAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 93595
Oy 649 aatgatgtgtaacagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 708
Db 93594 ACTTTTGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 93535
Oy 709 tgaatcgaagagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 765
Db 93534 ACATGCTTCAAGAAATTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 93475
Oy 766 gattcgaagctgcaatagagctgctgctgctgctgctgctgctgctgctgctgctgctgct 825
Db 93474 AATTGCTTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 93415
Oy 826 attgcatctgcaagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 860
Db 93414 TATATCTTTTGAATAATGCGCGGAGAAATGATGCTCG 93380

```

AUTHORS

Sugita, M.
Direct Submission
Submitted (08-MAR-1997) to the DBJ/EMBL/GenBank databases.
Masahito Sugita, Nagoya University, Center for Gene Research;
Chikusa-ku, Nagoya, Aichi 464-01, Japan (Tel: +81-52-789-3080,
Fax: +81-52-789-3081)

FEATURES

Location/Qualifiers
1..150613

/organism="Chlorella vulgaris"

/db_xref="taxon:3077"

/note="ORF1"

/codon_start=1

/protein_id="BAAS7837.1"

/db_xref="GI:2224353"

/translation="MKKLPFMSFSDSIFFLRNEKLIESKNSLFFFLRKKKDAFLP"

/complement(1936..2250)

/note="ORF104"

/codon_start=1

/protein_id="BAAS7838.1"

/db_xref="GI:2224354"

/translation="MAAASAESCSFEPFPAFSGALVKGSGVPIBOLLEHL"

/note="ORF105"

/codon_start=1

/protein_id="BAAS7839.1"

/db_xref="GI:2224355"

/translation="KMSHKEFEVLEKSEYKSRMARFRLKTSKTHDKITGNITV"

/complement(2889..3149)

/note="ORF86a"

/codon_start=1

/protein_id="BAAS7840.1"

/db_xref="GI:2224356"

/translation="MGITVFOFTVRSVSLRLNTINKTKAMOFVRRKVCWESSI"

/complement(3393..3551)

/note="ORF52a"

/codon_start=1

/protein_id="BAAS7841.1"

/db_xref="GI:2224357"

/translation="MRKVLNLKOTTFKSRSSASACCAIMCTGTFVSKSLQFSDV"

/complement(4095..5156)

/gene="psbA"

/complement(4095..5156)

/gene="psbA"

/codon_start=1

/product="photosystem II D1 protein"

/protein_id="BAAS7842.1"

/db_xref="GI:2224358"

/translation="MTAILRRSSASIMARCEWVSTENRLTYGMFGLMIPILLTA"

/complement(6212..6379)

/note="ORF55a"

/codon_start=1

/protein_id="BAAS7844.1"

/db_xref="GI:2224360"

CDS

Complete sequence of the green alga *Chlorella vulgaris* chloroplast genome reveals the existence of genes possibly involved in chloroplast division. (tentative)
Unpublished (1997)
3 (bases 1 to 150613)

```

RESULT 5
LOCUS AB001684 150613 bp DNA circular PLN 14-APR-2000
DEFINITION Chlorella vulgaris C-27 chloroplast DNA, complete sequence.
ACCESSION AB001684
VERSION AB001684.1 GI:2224352
KEYWORDS
SOURCE
ORGANISM Chlorella vulgaris chloroplast DNA.
REFERENCE
AUTHORS Chlorella vulgaris
Eukaryota: Viridiplantae: Chlorophyta: Trebouxiophyceae:
Chlorellales: Chlorellaceae: Chlorella.
1 (sites)
Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S.,
Ohta, T., Imamura, A., Yoshinaga, K., and Sugiyama, M.
TITLE Complete nucleotide sequence of the chloroplast genome from the
green alga Chlorella vulgaris: the existence of genes possibly
involved in chloroplast division
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (11), 5967-5972 (1997)
MEDLINE 97303241
REFERENCE 2 (sites)
Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S.,
Ohta, T., Imamura, A., Yoshinaga, K., and Sugiyama, M.
TITLE Complete sequence of the green alga Chlorella vulgaris chloroplast
genome reveals the existence of genes possibly involved in
chloroplast division. (tentative)
JOURNAL Unpublished (1997)
REFERENCE 3 (bases 1 to 150613)

```



```

/protein_id="BAA90610.1"
/db_xref="GI:6907083"
/translation="MAAALALAFVAAGALVAQADTPPYVEFDWDVSFLTASLGLIP
OKVIALIKQKPGVVMKLTITRYNNVNVNLISIDPELTIWDGLOONRNVIMQDQGLTTC
PIPCGMWITNBOVKQIGSEFVLTSLORAGSGEGITGNRAKRVISVPEPTDFDGI
TLFGDWTKASHDKRMLDDGKEGLMGKGVRYRNDLSVPLGIEHETIKVE

```

```

/note="hypothetical protein"
/codon_start=1
/protein_id="BAA90615.1"
/db_xref="GI:6907088"
/translation="MTLTRLRIQPGVSHVGRPPSEFIIDAKAFKPTQVIVYIMPTFGVDDQ

```



```

gene      /antlicodon=(pos:798..800,aa:Leu)
765..849
/gene="trnL(uaa)"
tRNA      /gene="trnQ(uug)"
983..1054
/note="codons recognized: CAA"
/gene="trnQ(uug)"
983..1054
/antlicodon=(pos:1015..1017,aa:Gln)
1102..1182
/gene="trnQ(uug)"
1102..1182
/note="trnL(uag)"
/gene="trnL(uag)"
/note="codons recognized: CUA"
/gene="trnL(uag)"
/antlicodon=(pos:1136..1138,aa:Leu)
1102..1182
/gene="trnL(uag)"
complement(1364..1468)
/gene="psbM"
complement(1364..1468)
/gene="psbM"
/codon_start=1
/product="M protein of photosystem II"
/protein_id="A054779.1"
/db_xref="GI:5880686"
/translation="MEVNIIGLITATLFIITPTSFLLILVKTAQON"
1618..1691
/gene="trnM(cau)"
/note="codon recognized: AUG"
/product="trnM-Met"
/antlicodon=(pos:1652..1654,aa:Met)
1618..1691
/gene="trnM(cau)"
1733..3862
/gene="trnI"
1733..3862
/gene="trnI"
/codon_start=1
/product="putative peptidoglycan synthetase"
/protein_id="A054780.1"
/db_xref="GI:5880687"
/translation="MIVNPMKYRQFRRLKRIHGGIYAFILISTMIKIALDPYKSR
PMKMHSGTPEYERERWVMTKLATVGLFAPRLSGIQFMYKTELKSVARQOIG
VNDROOKIILDHGDIATIDPATDILVHPRMCSISLERIAELSPILDLSSOVLN
RLDGDSGICLMHODITNTSAQIRLQVAGIELVHNPQVYPRKSGFESITIGYNTG
KVGGLSSLDMMKSTYQDVPCWMDGNGFLGIRPKQILFHOSALQTLIDGGLQ
KVSOLITNMRKFGAKRIALIMEHSGAIRLATSPEDNCGMPERFRCPTIT
DIFPGSTFRKVNIALALENGIQPTRIIDTGRKIRGDSWIGVCGFTDRSLDL
NGDUTSDVAISVKKDDEVDHEIEATASFGGLAMTPKLLQTLATLANGMAVT
PHLISKIVTLDFHILQSNMSEFSLQGWGVSRSQYHAKQPRYTHDILGVAVP
SLEGMVDKASIPNITERRRRLRRCOTCNVLGLLEQVLDQAPYSGRGLPGYAMV
KTGTAKASALGSGTSDVTSFVGIYPAKPKFVTLVIIIDEPDRFGFNTAVDT
QTLISEMIVQEOQPPSPVLSFEFRNA"
3998..5059
/gene="psbA"
3998..5059
/gene="psbA"
/codon_start=1
/product="D1 reaction-center protein of photosystem II"
/protein_id="A054781.1"
/db_xref="GI:5880688"
/translation="MTAILERRESTSVARPCDWTSTENRLYIGMGVLMPLLLTA
TSLVIGFIAPAPVIDIGTREPVSGLFENNTISGAILIPSSAIGIHFPIWEAASI
YPIGQGSFSDMGLGICSTENFMIVFOAHNHLIMPHMLGVAGVGGSLFSAMHGL
VTSSLIRETTENASNAGYKFGQDEETYNIAAHGYFGLRILFOYASFNNSRLHFLA
AMPVYCTWTLAGVSTMAFNNGFNFNOSVVDQGRVINTWADILIRANLIGEVNHER
NAHNPRLIASVDPAVQ"
5489..6964
/gene="atpB"
5489..6964
/gene="atpB"
gene      /codon_start=1
/product="CF1 beta subunit of ATP synthase"
/protein_id="A054782.1"
/db_xref="GI:5880689"
/translation="MSQAGVTKRQAKIYOIVGPNVDVAFQFQSPMNIYALIVKGRN
GAGQEVSVCEVOQLDGLVRAVMSAPDQMKRGCEVDTGKALSVAPGPTLGRIF
NVLGEVDMKRGVGNKTLPIHREAPAVYDLTKLSIFETGIKVDLLAPYRRGKIG
LFQAGVGTIVLMELINNIKAAHGVSVFGVGEKTRGNDLYMECESKYINKADV
KSKVALYVGOMNPPGARMRGVGLTMAEYFRVQVNDVLLFTIDNIFRVOAGSEV
SALLGRMPAGVQBPPLATEKSGLOERITTSQDSTTSQAVVPADLTPAPATF
AHLDAFTVLSRLAKGILPVDPLDSTSMQPMIILGDDHGTORVQOTLQRRKL
ODIITAIIIGIDELSEEDRMIVANARKIERLSQPFVVAEFGTAPGPKYVPLASIOGPK
6977..7378
/gene="atpE"
6977..7378
/gene="atpE"
/codon_start=1
/product="CF1 epsilon subunit of ATP synthase"
/protein_id="A054783.1"
/db_xref="GI:5880690"
/translation="MSIQVRILTPQVFLNVADEIILPTNGQGVLTNHPILITAI
DIGMIVRSSTQMSMDLGLLALVKNQVLIIVNEALGSDINAEAEETFLAKKA
LANSKTRKQDLENNLAFRRKRVQVATLVK"
complement(7622..8104)
/gene="petD"
complement(7622..8104)
/gene="petD"
/codon_start=1
/product="subunit IV of cytochrome b6/f complex"
/protein_id="A054784.1"
/db_xref="GI:5880691"
/translation="MSVTKKPDITDVLRAKLAKGNGHNYGEPAMPNDLLYFPPVI
LGLTSLTGLAVDPAIGEPANFATPLEIIPENYFPPFOLLRTVPKILGLILMA
AVRAGGLIVPFLESINKFQNPFRPVAIVTFLGVVAIMLGICATLPIDISTGLLF
complement(8132..8779)
/gene="petB"
complement(8132..8779)
/gene="petB"
/codon_start=1
/product="apocytochrome b6 of cytochrome b6/f complex"
/protein_id="A054785.1"
/db_xref="GI:5880692"
/translation="MSKITDNEERLEIQAIADITSKYPPHVNIFYCFGGITLTC
LIVQVAGTAMPFYIRPITVENFASVEYIMTNGFGLIRSHRMSAMMMMLTIHF
RYVLGTGKRRRELWVGVLAIVTSFEGVYSLPMDQVGMVAKIVTGPDPPIV
IGAPLVELLRQSVSGOSTLIRFYSLHFFVLPRLTAVMLMHLPLMKRQGISGFL"
complement(9045..9431)
/gene="rpl12"
complement(9045..9431)
/gene="rpl12"
/codon_start=1
/product="ribosomal protein L12"
/protein_id="A054786.1"
/db_xref="GI:5880693"
/translation="MSTITNQITEQLSMTLLEAELVSOIEETFGVDSAPAAAGVM
VMAAPSAPEVEQTEFTVINIDVPSKRIAVIKVVALTGLGKREKKMIESVPRI
KEGIGDEADQIIOOLEAGATVTK"
complement(9442..9846)
/gene="rps9"
complement(9442..9846)
/gene="rps9"
/codon_start=1
/product="ribosomal protein S9"
/protein_id="A054787.1"
/db_xref="GI:5880694"
/translation="MDHTSSVGLRTRKRCACVAQVRIIEGTGOLLINORPVPYLO
NPAYLLAOGALDVTETNRYDTIVTKTGGGGLTGOAEVILKGLANALCSLHIKHKAL
KPGGLITRNPLKRRKKYGLKRAAPQFSKR"
complement(9998..11485)
/gene="rpoA"
complement(9998..11485)

```


- AUTHORS Douglas, S.E. and Durnford, D.G.
 TITLE Nucleotide sequence of the genes for ribosomal protein S4 and tRNA(Arg) from the chlorophyll c-containing alga *Cryptomonas phi*
 JOURNAL Nucleic Acids Res. 18 (7): 1903 (1990)
 MEDLINE 90245597
 REFERENCE 4 (bases 18535 to 19351)
 AUTHORS Douglas, S.E. and Durnford, D.G.
 TITLE Sequence analysis of the plastid rDNA spacer region of the chlorophyll c-containing alga *Cryptomonas phi*
 JOURNAL DNA Seq. 1 (1): 55-62 (1990)
 MEDLINE 92119320
 REFERENCE 5 (bases 34539 to 35380)
 AUTHORS Reith, M. and Douglas, S.
 TITLE Localization of beta-phycoerythrin to the thylakoid lumen of *Cryptomonas phi* does not involve a signal peptide
 JOURNAL Plant Mol. Biol. 15 (4): 585-592 (1990)
 MEDLINE 91338697
 REFERENCE 6 (bases 110917 to 113854)
 AUTHORS Douglas, S.E.
 TITLE Unusual organization of a ribosomal protein operon in the plastid genome of *Cryptomonas phi*: evolutionary considerations
 JOURNAL Curr. Genet. 19 (4): 289-294 (1991)
 MEDLINE 91330343
 REFERENCE 7 (bases 40675 to 42376)
 AUTHORS Douglas, S.E. and Turner, S.
 TITLE Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor
 JOURNAL J. Mol. Evol. 33 (3): 267-273 (1991)
 MEDLINE 92099311
 REFERENCE 8 (bases 96129 to 98906)
 AUTHORS Wang, S.L. and Liu, X.Q.
 TITLE The plastid genome of *Cryptomonas phi* encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (23): 10783-10787 (1991)
 MEDLINE 92073372
 REFERENCE 9 (bases 106789 to 108216)
 AUTHORS Douglas, S.E.
 TITLE A secY homologue is found in the plastid genome of *Cryptomonas phi*
 JOURNAL FEBS Lett. 298 (1): 93-96 (1992)
 MEDLINE 92183838
 REFERENCE 10 (bases 42198 to 44153)
 AUTHORS Douglas, S.E. and Reith, M.E.
 TITLE A *hchl* homolog, encoding a subunit of Mg chelatase, is located on the plastid genomes of red and cryptomonad algae
 JOURNAL J. Mar. Biotechnol. 1, 135-141 (1993)
 REFERENCE 11 (bases 82327 to 84479)
 AUTHORS Douglas, S.E. and Murphy, C.A.
 TITLE Structural, transcriptional and phylogenetic analyses of the *atpB* gene cluster from the plastid of *Cryptomonas phi* (Cryptophyceae)
 JOURNAL J. Phycol. 30: 329-340 (1994)
 MEDLINE 912 (bases 98901 to 114602)
 REFERENCE 12 (bases 98901 to 114602)
 AUTHORS Wang, S.L., Liu, X.Q. and Douglas, S.E.
 TITLE The large ribosomal protein gene cluster of a cryptomonad plastid: gene organization, sequence and evolutionary implications
 JOURNAL Biochem. Mol. Biol. Int. 41 (5): 1035-1044 (1997)
 MEDLINE 97283757
 REFERENCE 13 (bases 61067 to 68605)
 AUTHORS Leitch, C.E.W., Kowalik, K.V. and Douglas, S.E.
 TITLE The *atpA* gene cluster of a cryptomonad, *Galliardia theta*: A piece in the puzzle of chloroplast genome development
 JOURNAL J. Phycol. (1998) In press
 REFERENCE 14 (bases 1 to 121524)
 AUTHORS Douglas, S.E. and Penny, S.L.
 TITLE The plastid genome of the cryptophyte alga, *Galliardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae
 JOURNAL J. Mol. Evol. 48 (2): 236-244 (1999)
 MEDLINE 99128221
 REFERENCE 15 (bases 1 to 121524)
 AUTHORS Douglas, S.E.
 TITLE Direct Submission
 JOURNAL Submitted (08-JAN-1998) Institute for Marine Biosciences, National Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1,

COMMENT On Sep 15, 1998 this sequence version replaced g1:11396 g1:11297
 g1:18103 g1:18281 g1:11383 g1:11407 g1:12539 g1:336730 g1:11300
 g1:398949 g1:2661180.
 FEATURES
 source
 location/Qualifiers
 1. 121524
 /organism="Galliardia theta"
 /organism="plastid:chloroplast"
 /db_xref="taxon:55529"
 /note="plastid"
 complement(5..86)
 /gene="trnY(gua)"
 /note="codon recognized: CAU"
 /product="tRNA-Tyr"
 /anticodon="(pos:50..52,aa:Tyr)
 complement(5..86)
 /gene="trnY(gua)"
 complement(181..573)
 /gene="rpl19"
 complement(181..573)
 /gene="rpl19"
 /codon_start=1
 /product="ribosomal protein l19"
 /protein_id="AAC35594.1"
 /db_xref="GI:3602933"
 /translation="MEIILHLEFVNKMSKNKLIQIVESNHNKTLPTQLQVGDVRLG
 VSIENKMERIQATEGVITSKSNGINLTTIRRVLGIGIERVYLHSPKLSIELI
 RSKIRRAKLYLTRSGKATRILOQRN"
 729..3188
 /gene="clpC"
 729..3188
 /gene="clpC"
 /codon_start=1
 /product="Clp protease ATP binding subunit"
 /protein_id="AAC35595.1"
 /db_xref="GI:3602934"
 /translation="MEERTEKAIKIVMLAOEBAARLGNHNVGTEQITLIGISGTGI
 AKVLKSMGVNLKDAVEVEKILGRSGSFVVEIPFTRARVLELSIEEROLGHP
 TGTTEHLLGLIREGEVAVARLENALDITRVQIVLIDPFAVLSATKQGTNP
 TLEFGSNTLTKAAGKLDPEVIRGKEIEIRVQIIGRTKNNPILIGBPQKTAIAE
 GLAQRINNDVDPIIDEDRVVPLDGLIVAGIKYGEFEERLKKITIDIRVANNVIL
 IDEVHTLIGAGAEGAIIDANTIKPALKRGMOCIGATLLEERKIDSLERKQ
 LIDEGASRWRLNNSQDPPARELDELRLKILITKQDEAVSODFETAGOLRDEMETIL
 PWVGEPSVEETIETILGRDREYKHHKTIIDSLAFAADQYIDRLPDAID
 OLAIAHSKKDGEENTREVSVTEEDIAIVAMGIVPNKMTRESKILOMETIKL
 GRIGODEAVVAVSKAIRARVGLKPNPVPVASFISGFTGVKRELKALASVYFGS
 EAMVRLDSEVEMERTVSKLIGSPGVYNGDQLTESVRRPPTVYVPEIKKH
 PDPYNTLITLIEDGRITDYSKGRVDFKNTLLITSNVSKVIEKGGGIGVLESDQI
 ESOYGRKALVNEELKQYRPEPLNRLDEIIVFROLTKDEVEIEIMLKEVFTNIS
 KGIOLEVTARFKTHLIDEGNPIYGARVPLRAVMRLLEDTLSEELAEKIKRGDAV
 DVDDGKRVLLDSEKLELVAN"
 3316..3388
 /gene="trnF(gaa)"
 /note="codon recognized: CUU"
 /product="tRNA-Phe"
 /anticodon="(pos:3349..3351,aa:Phe)
 3316..3388
 /gene="trnF(gaa)"
 complement(3402..5201)
 /gene="dnab"
 complement(3402..5201)
 /gene="dnab"
 /codon_start=1
 /translation="pos:complement(5199..5201),aa:Met)
 /product="replication helicase subunit"
 /protein_id="AAC35596.1"
 /db_xref="GI:3602935"
 /translation="MKVTRFYNLAKAEKHLIIQILVEKDFLOITKIDTQVYFROSHO
 ILIYSINILHTRKVLINPLNLSLISELSEVETVINYLTIFNDNRPLININDSKL
 VLDNFTIRELONSCAKIVSLTFNPNLSLSEVETVINYLTIFNDNRPLININDSKL
 LLEETLEIDKTNRESTRVLTGPFDDHLIVGLQSDLLIIGRSPSMGTAFMLSLVRN
 VADIOSFPIVPSLEMSKQILYRLISMETINATSRLEGNSISEWILRAMVTILS
 NNTIYLDENNDIVDIOKSLASLOQIYDGLIVIDVLQLOKREKSSQRIELSYI

[illegible]

cds.
 AF149810 GI:6002787
 AF149810.1 GI:6002787
 ORIGIN
 ORGANISM
 SOURCE
 ORYZA SATIVA
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
 1 (bases 1 to 673)
 Han, F., Kilian, A., Chen, J. P., Kudrna, D., Steffenson, B.,
 Yamamoto, K., Matsumoto, T., Sasaki, T. and Kleinof, A.
 Sequence analysis of a rice BAC covering the syntenous barley Rpg1
 region (1999) In press
 2 (bases 1 to 673)
 Han, F., Kilian, A., Chen, J. P., Kudrna, D., Steffenson, B.,
 Yamamoto, K., Matsumoto, T., Sasaki, T. and Kleinof, A.
 Direct Submission
 Submitted (12-MAY-1999) Crop and Soil Sciences, Washington State
 University, Pullman, WA 99164, USA
 FEATURES
 source
 1
 location/Qualifiers
 673
 /organism="Oryza sativa"
 /cultivar="IR-BB21"
 /db_xref="taxon:4530"
 /chromosome="6"
 /note="covering the syntenous barley Rpg1 region"
 complement(1..537)
 /codon_start=1
 /product="cell division inhibitor MIND homolog"
 /protein_id="AF00142.1"
 /db_xref="GI:6002787"
 /translation="MAFAPRLPLSPRCPPASSPARHGRAPETLGPTRVYVTS
 KGVGVTTTANLAASLARSLSAVADADAGLRNLDLGLERVYLTADAVIAGGC
 RUDQALVRRHALDLGLCLSKPRSKPLAFGSKTLTWADALRRANPFIILIDCP
 AGSAPPEPSTYDDGCM"
 BASE COUNT
 115 a 160 c 295 g 103 t
 ORIGIN
 160 c 295 g 103 t

AP001517 BA000004
 AF001517.1 GI:10175500
 AF001517.1 GI:10175500
 ORIGIN
 ORGANISM
 SOURCE
 BACILLUS HALODURANS DNA
 Bacteria; Firmicutes; Bacillus/Clostridium group;
 Bacillus/Staphylococcus group; Bacillus.
 1 (sites)
 Takami, H., Nakasone, K., Hirama, C., Takaki, Y., Masui, N., Fuji, F.,
 Nakamura, Y. and Inoue, A.
 An improved physical and genetic map of the genome of alkaliphilic
 Bacillus sp. C-125
 Extremophiles 3 (1), 21-28 (1999)
 2 (sites)
 Takami, H., Nakasone, K., Ogasawara, N., Hirama, C., Nakamura, Y.,
 Masui, N., Fuji, F., Takaki, Y., Inoue, A. and Horikoshi, K.
 Sequencing of three lambda clones from the genome of alkaliphilic
 Bacillus sp. strain C-125
 Extremophiles 3 (1), 29-34 (1999)
 3 (sites)
 Takami, H., Takaki, Y., Nakasone, K., Hirama, C., Inoue, A. and
 Horikoshi, K.
 Sequence analysis of a 32-kb region including the major ribosomal
 protein gene clusters from alkaliphilic Bacillus sp. strain C-125
 Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)
 4 (sites)
 Takami, H., Masui, N., Nakasone, K. and Horikoshi, K.
 Replication origin region of the chromosome of alkaliphilic
 Bacillus halodurans C-125
 Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)
 5 (sites)
 Takami, H., Takaki, Y., Nakasone, K., Sakiyama, T., Maeno, G.,
 Sasaki, R., Hirama, C., Fuji, F. and Masui, N.
 Genetic analysis of the chromosome of alkaliphilic Bacillus
 halodurans C-125
 Extremophiles 3 (3), 227-233 (1999)
 6 (sites)
 Takami, H.
 Genome analysis of facultatively alkaliphilic Bacillus halodurans
 C-125
 (in) Extremophiles in deep-sea environments (Ed.):
 HORIKOSHI, K. TSUJII:
 : 249-284; Springer-Verlag (1999)
 7 (sites)
 Takami, H. and Horikoshi, K.
 Reidentification of facultatively alkaliphilic Bacillus sp. C-125
 to Bacillus halodurans
 Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
 8 (sites)
 Takami, H. and Horikoshi, K.
 Analysis of the genome of an alkaliphilic Bacillus strain from an
 industrial point of view
 Extremophiles 4 (2), 99-108 (2000)
 9 (bases 1 to 294250)
 Takami, H., Nakasone, K. and Takaki, Y.
 Direct Submission
 Submitted (22-MAR-2000) to the DDBJ/EMBL/GenBank databases. Hideto
 Takami, Japan Marine Science and Technology Center, Deep-sea
 Microorganisms Research Group, 2-15 Natsushima, Yokosuka, Kanagawa
 237-0061, Japan (E-mail: takami@jstec.go.jp)
 URL: http://www.jstec.go.jp/jstec/e/plo/DEPSTAR/Research.html,
 Tel: 81-468-67-3895, Fax: 81-468-66-6364)
 location/Qualifiers
 1..294250
 /organism="Bacillus halodurans"
 /db_xref="taxon:86665"
 /note="alkaliphile"

gene 236..823
/gene="BH2879"
CDS 236..823
/gene="BH2879"
/codon_start=1
/transl_table=11
/product="spore cortex-lytic enzyme"
/protein_id="BAB06598.1"
/db_xref="GI:10175501"
/translation="MMKPVLLPSCILSLPFLFNKAASHSKHNOVSGPILYLLSBOY
GVPMELIKRINERSSNTIYRGEOLITPAVATASRELLARIVAEAGEPAGKVA
VVVLRNVDHSPEDPNTVDVINEVSSGVYFSPONQRICORPANDEKRAVDLORRG
OGNGSLYFNPKATISPVSTRQOITIGNHLPK"
955..1605
/gene="BH2880"
955..1605
/gene="BH2880"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB06599.1"
/db_xref="GI:10175502"
/translation="MRRISRTSSRTGRTAVYVNPGLTSDSVKRVPEPPPIERTANET
AKPSNNVLSYERKQYKLEKQVSEFYHEKELYHTLQAGNKELISQCTLLC
KYNELTKSIRKLQDFOTSTMDIQAHLIOHOKLELELGLNDNGTLTNSDVFRC
LSTSDPTMAVLKRFKSLKLYHTISLOLPNSHONSPYDPCEPLKGTIEKG"
1694..2977
/gene="BH2881"
1694..2977
/gene="BH2881"
/note="unknown conserved protein in others"
/codon_start=1
/transl_table=11
/protein_id="BAB06600.1"
/db_xref="GI:10175503"
/translation="MSEHSEHSYSDNLIDLMHTLTNDHLQLLKIGEMAAKLRML
FLVGTVRMDRGVGGDLVIEGDALFQSNVANYGKVKHHEPPATVWGAEN
LKDIIVASRAESVAKPGALPTIRHSHITDRLARDEFSINAMAILHPSYGLDPFH
GRHDLNHLILHLOSRTIDPDLRLRGVRSRNRYRFOKTAALATOPALTNAL
ANVSPTRIVELKLCHEFTIDPVSSSKIEDILAVQALILGFSSSATHLRLQLEON
GEPLHMFQAIATVGGLEDNMKASLVPPATITAMEQRFIONIEDICKRLTNMTRESTDYL
HKOLYVPEEPLREFVALSSGEMQVLDLYLHOKQLOPLLTGHDMLGKPSPLFK
ECLLHECEOLKGTLENKODALQFAREFFNHKOPL"
complement(3074..4309)
/gene="BH2882"
complement(3074..4309)
/gene="BH2882"
/codon_start=1
/transl_table=11
/product="3-oxoacyl-(acyl-carrier protein) synthase"
/protein_id="BAB06601.1"
/db_xref="GI:10175504"
/translation="MNRHYVVGIGAVPPLTGNVETWEKATAGOSGVGLTRVEAAD
FPMKVAAYEKEDPTTHIDRKARKMDRTOPAVASALMALDADLEIEDIAPRGV
WIGSIGMETYENDRPMFTEKGRVRSFPYPMATMASGOVSTITGARIGNSGV
TACASGANSIGDAFRVIOQGDADVMITGGAELVITTMAYAGFSMAAITSDDPNTAS
RPDANRDVGEESGILILELSAEKRGAKIYAEIVGATGDAIVTODAPBGE
GAARAKQAIEDAGLPEDIOYMANHSTGYNDKXETLAQVDFDHYNNALISSTK
SMTGHLGAAGVAEALFESYKALEEGLIPTINVEPDDCDLDYVNEARKOGVNAV
SNSLGGGHNATVEKKY"
complement(4347..5285)
/gene="BH2883"
complement(4347..5285)
/gene="BH2883"
/codon_start=1
/transl_table=11
/product="3-oxoacyl-(acyl-carrier protein) synthase"
/protein_id="BAB06602.1"
/db_xref="GI:10175505"
/translation="MTRAGILIGSYVGEHVYTNKDFEKRLIDSDEIRTRTGIEER
FAPNDVTSMAEYASAKAGVSAEIDLILVATVTPDMPPTVSTIVOOHLCK
KAAAMDISACAGETIYGLATGOOFIENGGRKHYLVIGVERLSKITDMDTNTAVLEED
GAGAAVAGVASEDRLIAYELGADGEGAMHINQGEYIQMNGREVEKFAVRKMGESAL

SVLEKAGLSKEDVDLFIHQANIRIMEASRERLELPVEKMTTTRKNGTSSASIPMA
MYDELKDGKIKDGDLLVYFGAGLWMSLALRMGR"
5508..5807
/gene="BH2884"
5508..5807
/gene="BH2884"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB06603.1"
/db_xref="GI:10175506"
/translation="MSNOERDFFTLMLGRPESTDVAEENPQESTGVEETKE
GETKQETNAPALPLTFROMEHIFRLAOSLGLKSLSPYKTIQOLLSQKKD"
complement(5903..6157)
/gene="BH2885"
complement(5903..6157)
/gene="BH2885"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB06604.1"
/db_xref="GI:10175507"
/translation="MNYTHGPYROYIGTGERFPAVLPLAGLAVPLLEGKRP
YGGYGPAYGSGPNPBNYGNPGRNPGPYPGPPSOPPFGYK"
complement(6226..6555)
/gene="BH2886"
complement(6226..6555)
/gene="BH2886"
/note="unknown"
/codon_start=1
/transl_table=11
/protein_id="BAB06605.1"
/db_xref="GI:10175508"
/translation="MNKEDALRSFYEVKKKKQOLKDKRFEETIOTLPYIELFRQTD
VAEPQVFSYVIAQLRTGEFEGFLRTVEDIKGEMELTEAEKAVKLEGLFHLVDLAQL
AASDKQL"
complement(6636..7421)
/gene="BH2887"
complement(6636..7421)
/gene="BH2887"
/note="unknown conserved protein in others"
/codon_start=1
/transl_table=11
/protein_id="BAB06606.1"
/db_xref="GI:10175509"
/translation="MGIGSSNRQTDMMNAKLYDEHRRFVSAYGEDLIOMLAPKECV
IDLGGCTDLITQIHLQDSRVIGVSVSMIRQAGKPPHLPFOVAEATDLSERFD
ALFSAVNLHMIKDAEALVITYRSLLKPGSRPYAFRGGKGNVETIYNTLADTMSQDSR
WSELELPWFEPFISIGYTSIMEOVGFRTVHALHPNRPITLLEDDGIRHMLAMAEVYLE
GKSETERAEIVFEAEVDALFDGNTMMADYKRIRIVGMKOL"
complement(7607..7843)
/gene="BH2888"
complement(7607..7843)

Query Match 7.28; Score 84.8; DB 2; Length 294250;
Best Local Similarity 50.08; Pred. No. 4.4e-09;
Matches 252; Conservative 0; Mismatches 237; Indels 15; Gaps 1;
OY 238 aatcgacgcgacgttgctcaacccgctctctctcgcgcgcgaaacgcgcgt 297
DB 145487 AGTCGACACGACATTTGGCTTGCATATTAGATGGTGATGGTTAGAAAATCCGAT 145428
OY 298 caattacacgcttgtaagtlctcaacgagcattcgaacccgacccaagccctagctcg 357
DB 145427 TATTTATGACCTTGTGATGATGAGGAGGCGCTGTCGGCTAAAGCAGCGGCTCATCAA 145368
OY 358 tgataaacgcctgcaattcgaatgcttgtaattccaacactagylcaaaatlgcc 417
DB 145367 AGATTAACGCTTGTGAATTTGTTACCTGCTCGGAAAGAAAGATAAGTCTGC 145308
OY 418 tttagatttgagggaagaagcttagcttgctgagtcattaaagaatagcaagaag 477
DB 145307 AGTAACCCCTTAACAAA-----TGAAGGAAATCGTCGAGGAGCTTAAACA 145263

Qy 478 ttgcccgaattatatactatagatgtctcaggaattatgatgcgaggttaacacgc 537
 Db 145262 AGAATATGACTATCTTATCATGATTCGCGGTGGGATCGAACAAGGCTTTAAAAATGC 145203
 Qy 538 catcacacgcgttaacgaagacgcgtattatcaacacactgatactatcgttagaga 597
 Db 145202 GGTCCGGGGGGGATGAAAGCATGCTGTACAACTCCTGAATTTCTTCAGTCCGTGA 145143
 Qy 598 tgcagatagattacagcgtctgtaagtgatggaattaggaattataaattgattgt 657
 Db 145142 TGTCTGATCGAATTCATTTGGCCTTCTCGACAAGAGAGGTTGAAGCACCCTCTCGTCCT 145083
 Qy 658 gaacagagttagaactgttttgataaagggtgaagataatgctatgtcttgatgtca 717
 Db 145082 CAATCGAATCCGTGCGCCATATGATGATAAAGCAGACAGATGCTGATGTGAGAGATTTGT 145023
 Qy 718 agagaattcgtggaattgtatgtatgt 741
 Db 145022 TTCCATCTTACCATGCACTATTT 144999

RESULT 15
 AEO01931/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS

AEO01931 16486 bp DNA BCT 22-NOV-1999
 Deinococcus radiodurans R1 section 68 of 229 of the complete
 chromosome 1.
 AEO01931 AEO00513
 AEO01931.1 GI:6458461

Deinococcus radiodurans.
 Deinococcus radiodurans.
 Bacteria: Thermus/Deinococcus group: Deinococcales; Deinococcus.
 1 (bases 1 to 16486)
 White, O., Eissen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D.,
 Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L.,
 Morfitt, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
 Vamathevan, J.J., Lam, P., McDonald, L., Uitterback, T., Zalewski, C.,
 Makarova, K.S., Aravind, L., Daly, M.J., Fraser, C.M. et al.
 Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans R1
 Science 286 (5444), 1571-1577 (1999)
 2 (bases 1 to 16486)

TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS

White, O., Eissen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D.,
 Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L.,
 Morfitt, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
 Vamathevan, J.J., Lam, P., McDonald, L., Uitterback, T., Zalewski, C.,
 Makarova, K.S., Aravind, L., Daly, M.J., Minton, K.W.,
 Fleischmann, R.D., Ketchum, K.A., Nelson, K.E., Salzberg, S.,
 Smith, H.O., Venter, J.C. and Fraser, C.M.
 Direct Submission
 Submitted (08-NOV-1999) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 Location/Qualifiers
 1. 16486
 /organism="Deinococcus radiodurans"
 /strain="R1"
 /db_xref="taxon:1299"
 /chromosome="1"
 /complement(140..427)
 /gene="DR0751"
 /note="similar to PID:1001780 percent identity: 53.57;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0330.1"
 /db_xref="GI:6458462"
 /translation="MGWFKGRSKETLKRLLEVLAVYRAKIAQKVEALRNDLLEY
 VQKTFSGNSVVEVDGDKVLMASISIEASISTEEESAAEDKSDKPSA"
 complement(427..1257)

TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS

White, O., Eissen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D.,
 Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L.,
 Morfitt, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M.,
 Vamathevan, J.J., Lam, P., McDonald, L., Uitterback, T., Zalewski, C.,
 Makarova, K.S., Aravind, L., Daly, M.J., Minton, K.W.,
 Fleischmann, R.D., Ketchum, K.A., Nelson, K.E., Salzberg, S.,
 Smith, H.O., Venter, J.C. and Fraser, C.M.
 Direct Submission
 Submitted (08-NOV-1999) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 Location/Qualifiers
 1. 16486
 /organism="Deinococcus radiodurans"
 /strain="R1"
 /db_xref="taxon:1299"
 /chromosome="1"
 /complement(140..427)
 /gene="DR0751"
 /note="similar to PID:1001780 percent identity: 53.57;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0330.1"
 /db_xref="GI:6458462"
 /translation="MGWFKGRSKETLKRLLEVLAVYRAKIAQKVEALRNDLLEY
 VQKTFSGNSVVEVDGDKVLMASISIEASISTEEESAAEDKSDKPSA"
 complement(427..1257)

gene
 CDS

/gene="DR0752"
 /complement(427..1257)
 /note="DR0752"
 /note="similar to PID:1001780 PID:1001783 percent
 identity: 76.89; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="septum site-determining protein"
 /protein_id="AAI0331.1"
 /db_xref="GI:6458463"
 /translation="MPCYAPASMDAKYIVYVSGKGVGKTTTANIGAAALRIQEKVY
 VIDYDGIARLDIVYVGLSEGRVYFPLVLEDEKCRNMNALIRDKRENIHLLPASQTRD
 KDADPEVFEKSVYVGLSEEGEDRYLIDSPAGIESGFETAPAPAEALVYVNEVSSV
 RDADRITGLLEAODIYTERILVYVNRILRPYVMSGMLSIDQVVDILGKPIGIVPEDEG
 IIVSTNWGEPAVILCKTKACDAFMATQRIQODVPPPKLTREEKGIWAAIRLFGCS
 1275..2063
 /gene="DR0753"
 1275..2063
 /gene="DR0753"
 /note="similar to GB:AL009126 percent identity: 50.23;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0332.1"
 /db_xref="GI:6458464"
 /translation="MGVEAGCGVACVPGGDSQPECRRRAGRLPGSMTRILGNNONR
 SVLIIGTLVAALALIALPAVRKSTSTTGEGQTFYANLPYAGANAVNVLNNGNOR
 KCPNCKSPFEYVAPLRLKIVGTGKRWMSLYVPLADRLPEDDSKYAQAARCYVYAO
 GKDAENYTKETLEFRAGPEYVWATKSKELASLIDIDOKAFPTCLDNDETAQVE
 2074..2535
 /gene="DR0754"
 2074..2535
 /gene="DR0754"
 /note="similar to GB:AL009126 percent identity: 65.89;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0333.1"
 /db_xref="GI:6458465"
 /translation="MNRDTRTLYLAWLVAATIGSLYFSEIRHFNPCPLCAORIFWY
 PLAVIIGIAAEVGGHVRVYPLAALCTGATITONLEWGEVOSIACIVYNAACN
 2683..3183
 /gene="DR0755"
 2683..3183
 /gene="DR0755"
 /note="similar to SP:P30529 percent identity: 75.44;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="ribosomal protein L19"
 /protein_id="AAI0334.1"
 /db_xref="GI:6458466"
 /translation="MOTHIKINRGELLGRIEDDHVROLPDFRPGDTVRVDTVREGRN
 TRSOAPEGVYVAINIGSGSRKSTFVRKISGEVEYVFPAPSLVNVQTVTVRGGKRRRA
 KLYVLELNGKRAARKSDRSRVKDAARAOQOKANASASQMAAQAADVIVISAEVVA
 PETQGS"
 complement(3254..3826)
 /gene="DR0756"
 complement(3254..3826)
 /gene="DR0756"
 /note="similar to GP:3449262 percent identity: 52.70;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0335.1"
 /db_xref="GI:6458467"

gene
 CDS

/gene="DR0752"
 /complement(427..1257)
 /note="DR0752"
 /note="similar to PID:1001780 PID:1001783 percent
 identity: 76.89; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="septum site-determining protein"
 /protein_id="AAI0331.1"
 /db_xref="GI:6458463"
 /translation="MPCYAPASMDAKYIVYVSGKGVGKTTTANIGAAALRIQEKVY
 VIDYDGIARLDIVYVGLSEGRVYFPLVLEDEKCRNMNALIRDKRENIHLLPASQTRD
 KDADPEVFEKSVYVGLSEEGEDRYLIDSPAGIESGFETAPAPAEALVYVNEVSSV
 RDADRITGLLEAODIYTERILVYVNRILRPYVMSGMLSIDQVVDILGKPIGIVPEDEG
 IIVSTNWGEPAVILCKTKACDAFMATQRIQODVPPPKLTREEKGIWAAIRLFGCS
 1275..2063
 /gene="DR0753"
 1275..2063
 /gene="DR0753"
 /note="similar to GB:AL009126 percent identity: 50.23;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0332.1"
 /db_xref="GI:6458464"
 /translation="MGVEAGCGVACVPGGDSQPECRRRAGRLPGSMTRILGNNONR
 SVLIIGTLVAALALIALPAVRKSTSTTGEGQTFYANLPYAGANAVNVLNNGNOR
 KCPNCKSPFEYVAPLRLKIVGTGKRWMSLYVPLADRLPEDDSKYAQAARCYVYAO
 GKDAENYTKETLEFRAGPEYVWATKSKELASLIDIDOKAFPTCLDNDETAQVE
 2074..2535
 /gene="DR0754"
 2074..2535
 /gene="DR0754"
 /note="similar to GB:AL009126 percent identity: 65.89;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0333.1"
 /db_xref="GI:6458465"
 /translation="MNRDTRTLYLAWLVAATIGSLYFSEIRHFNPCPLCAORIFWY
 PLAVIIGIAAEVGGHVRVYPLAALCTGATITONLEWGEVOSIACIVYNAACN
 2683..3183
 /gene="DR0755"
 2683..3183
 /gene="DR0755"
 /note="similar to SP:P30529 percent identity: 75.44;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="ribosomal protein L19"
 /protein_id="AAI0334.1"
 /db_xref="GI:6458466"
 /translation="MOTHIKINRGELLGRIEDDHVROLPDFRPGDTVRVDTVREGRN
 TRSOAPEGVYVAINIGSGSRKSTFVRKISGEVEYVFPAPSLVNVQTVTVRGGKRRRA
 KLYVLELNGKRAARKSDRSRVKDAARAOQOKANASASQMAAQAADVIVISAEVVA
 PETQGS"
 complement(3254..3826)
 /gene="DR0756"
 complement(3254..3826)
 /gene="DR0756"
 /note="similar to GP:3449262 percent identity: 52.70;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAI0335.1"
 /db_xref="GI:6458467"

```

/translation="MTVSAPVPLSVTLALDTATPPLTAVSPGCELPRNOVGRA
SARRIADETRALGAGLPEAOQIVIGTGPSTGVRGASVYALRGAGVAVRGV
STLESVIRGSGPOAVSLDARKGVSAVYEVSGVYDVQVARRPLAEFAELSL
PWRDPPPSGLALLRAGLHAGAKDWELATL"
/gene="DR0757"
3862..5097
3862..5097
/gene="DR0757"
PID:2293267 GB:AL009126 percent identity: 66.76;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="citrinate synthase"
/protein_id="AAFI0336.1"
/db_xref="GI:6458468"
/translation="MCPARSHLFVNPACTPDRPFPAWKKGCHYPLLTMSNAKLEG
VLTFSKTEFTNGSRGILTHLGIPIQEWAKSTFELSLALDAKLPAEELAKDAE
LKANRATFDOLVGTIRDMPKGVHPMAQALTVASYGLDPOMEDITPEARRAISTRM
AOFSTIIAIAINRQSGODIVAPRADLTUAGNFMLTGNEPTPEQARLELDIALVHAD
HGMNSTFTALSTSLSDPMSCWYSAICAKGLPLHGANEAVMTMLDEIGTVDAEA
YTGKLDNKEKIMVGHRYKYFDRSRVLRDYARHVANKBEKSNVYOLLEAIKIIY
DMGAKGITPNVDYSGTVSDLGIKKEYFTPLALRLSGMCASVIEYSQNRLLRP
DAEYTGARDOHYVDIKDRQ"
complement(5166..5993)
/gene="DR0758"
complement(5166..5993)
/gene="DR0758"
/note="similar to GB:AE000657 percent identity: 53.73;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFI0337.1"
/db_xref="GI:6458469"
/translation="MRYVICGGGTDEDDWAPGLIYGVKKDAFLITYAHNFDAVELN
SSFYAIPGLAEFGWVRKSSGVWRFAYKILKFTVHVAQPDITDPRMLQSPEDREG
MMGYLAQFYSFHRITANKRYLALLTERPRAGHLEIARHSAKDEVEBGEFEI
IWSPPDYVPGMPEQVHNVGVDGYRLHGRNRETYWEGSAGEHRDHYLNRAEME
WAKTIAVAGDISELYVFQNTTGHAKLNIPMLREALNANGVYVQTPAHDGGRLL"
6162..7217
6162..7217
/gene="DR0759"
6162..7217
/gene="DR0759"
/note="similar to PID:1550711 GB:AL123456 percent
identity: 58.59; identified by sequence similarity:
putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFI0338.1"
/db_xref="GI:6458470"
/translation="MTRSYSHPPYADPRAPHRETSPOELTAULTPGARPDVRFENYR
PPEFPSSQARDOLAEEVYRQAGKASGGGFWPFRKRAPREGGYLDGGEGVKTILL
ASAYHAAGKGGSGGSDVAFMSEFDMLYLIGALGMPQAIETERNYKLLIDEFELDD
GNTHMANFTLGLMPSGVSVATSNTEPGALAGRNADFOROIGIADRCGPHRID
GPDYRQRCGTEPQPLSEAEFAAMLARONPOTATLHAEILNRLLSVHRSRRRVLEK
VSCVATNLEPMPEQGEALRFVHFVDKVVYDLGISAFTGVLNALPDDIYRMNGYAKK
YSRALRSLSEMLREARG"
7245..8270
7245..8270
/gene="DR0760"
7245..8270
/gene="DR0760"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFI0343.1"
/db_xref="GI:6458475"
/translation="MFRRDPLMAQWRAVARILLEDDQARALLGIIVROTLLRRGGVLGV
ARGDERQLTFGVGVENGVEELASVTKPFTAAIASALVRGRLMNNAPLALGGPLLR
LPRATVYALATHTNAGLPPQPARALITFTFRADPYGMSPADVLASARRWANGQAG
RFGISNLGAGVIALGLHAAGEETSAAGYKRALRLRVTFPLGPGVGLTPADVPPY
```

```

gene
GILGCAVYGFAPLLAGAGGLFGSAELLHFGAEHLHSGAAGQHWROAFAFFGLPYAG
AAGGWRQSGSTVWHHDGARTRALGSPRSGAVVTLVVGAVPLVGVAVGVPDLLLG
LIGATDR"
complement(8372..9298)
Query Match 6.5%; Score 77.2; DB 1; Length 16486;
Best Local Similarity 48.3%; Pred. No. 2.1e-07;
Matches 254; Conservative 0; Mismatches 263; Indels 9; Gaps 1;
Oy 204 cagagctgcgcggagacacactccagcagatcgtcgcacatgacgcgacgttgctcagta 263
DB 1157 CGGCCCTGCGCCGACGCTGGTGAAGAAAGTGTGTGATTCACCTGACGAGTGGGTTGGCGGA 1098
Oy 264 acctgattcttcctcgcgtcgcgaacacgcgtcaattacacgcgtcgttgaagtctca 323
DB 1097 TTTTGGACCTGGTGTATGGCTTATAGATCCCGGTGTGTTCATTTGGTGGACGTGCTGG 1038
Oy 324 acgcgattgcagactcgacccaagccctagttcgtgataaagcgtcgtaaatctcgaat 383
DB 1037 AAGGCAAGTGGCCGATGAATCAAGCACTGATCCGCAAGACGCGGTGGAGAACCTGCACC 978
Oy 384 tgccttgatttcaaaaacctaggtcaaatctgaccttaggatttggggaaacgtttag 443
DB 977 TGTGCGCCGCTGCGACACCCGCGACAGGACGCCCTGGACCCGAGGCTTTCAGAGAG 918
Oy 444 ttgcgttgatgcaataaagataagcagaagaagttgcccggatttttaactatagat 503
DB 917 TCGTCAGGCGCTGCT-----GGAGAGAGAGGGCTTTCGACCGGCTGCTGACGT 867
Oy 504 gtccctgagatgatgagcgggttcaataccgcatatacacgcgctaaagcagat 563
DB 866 CGCCCGCGGATCGAGTGGGCTTCGCGACCGCGCGCCCGCGAGGGCCCGCTGG 807
Oy 564 tagttacacacactgatatctatcgtcatgtgagatgacagatgataaggttcaagcttgctg 623
DB 806 TGTGCTGTCAGACCCGAAAGTGTGTCGTGTCGCGCGACGCGACCGCATCATCGGCTGCTCG 747
Oy 624 aatgtgaggaattagggatatataaattgatttgaacagagttgaaactgatttgataa 683
DB 746 AAGCGCAGCAGATTCACCGAGATTTCGGCTGCTGTCACCGCTCGCTCCCAAGATGTGTG 687
Oy 684 gagggtgaagatgatatgltcagttcgttatgttcaagagatgtgttg 729
DB 686 CGAGCGGCAACATGCTCTCATCGACGACATGTGGACATCTCTGGG 641
```

Search completed: January 8, 2001, 09:29:36
Job time: 21190 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2001, 09:29:10 ; Search time 118.73 Seconds

(without alignments)
3739.858 Million cell updates/sec

Title: US-09-553-431-3

Perfect score: 1182

Sequence: 1 aagcttgatcgcacatcc.....tcaaaaaaaaaaaaaa 1182

Scoring table:

IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N-Geneseq_36:*

- 1: /cgn2_2/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /cgn2_2/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /cgn2_2/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /cgn2_2/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /cgn2_2/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /cgn2_2/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /cgn2_2/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /cgn2_2/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /cgn2_2/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /cgn2_2/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /cgn2_2/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /cgn2_2/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /cgn2_2/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /cgn2_2/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /cgn2_2/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /cgn2_2/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /cgn2_2/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /cgn2_2/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /cgn2_2/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /cgn2_2/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /cgn2_2/gcgdata/geneseq/geneseq/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	42	3.6	1664976	19 V21209	
2	40.8	3.5	933	21 A08365	Methanococcus jann
3	40	3.4	1049	19 V43610	A. thaliana trans
4	40	3.4	2885	20 200361	Human secreted pro
5	39.6	3.4	1223	20 X25150	Nucleotide sequenc
6	39.6	3.4	6761	20 X20517	Soybean isoflavone
7	39.4	3.3	1217	21 X20146	Polynucleotide seq
8	39.2	3.3	2286	20 X24385	Capicpum annuum (r
9	38.9	3.3	2328	20 X99549	Arabidopsis thaliana
10	38.6	3.3	1264	19 V12767	Nucleic acid sequen
11	38.6	3.3	5852	12 O11710	Metal-regulated tr
C 12	38.6	3.3	9789	17 T41852	Dicystostelium plas
					CDNA encoding Plas

13	38.2	3.2	1713	18 T85993	Maize 5-enolpyruv
14	38.2	3.2	1713	18 T86000	Maize 5-enolpyruv
15	38.2	3.2	1713	19 V19277	Maize 5-enolpyruv
16	37.6	3.2	2064	20 X37255	Human gene encodin
C 17	37.6	3.2	1476	20 X13679	Enterococcus faeca
18	37.2	3.1	2781	13 Q13239	HSF CDNA sequence.
19	37.2	3.1	2781	13 Q25712	Sequence of Drosop
20	37.2	3.1	2115	16 N91043	Decay accelerating
21	37.2	3.1	2115	16 Q79963	Decay accelerating
C 22	36.8	3.1	12381	14 Q54144	Sequence of plasm
23	36.8	3.1	12381	21 Z58381	Streptomyces averm
24	36.4	3.1	574	21 Z51738	Catalpa speciosa t
25	36.4	3.1	1897	20 X88528	Low density lipopr
26	36.4	3.1	1897	20 T32577	Bovine LDL recepto
27	36.4	3.1	1906	17 T32577	Low density lipopr
28	36.2	3.1	1205	18 T92824	Flea saliva protei
29	36.2	3.1	1205	19 V73413	Flea saliva protei
C 30	36.2	3.1	1205	19 V73413	Flea saliva protei
31	36.2	3.1	1559	21 A26377	Human secreted pro
C 32	36.2	3.0	904	20 X61467	DNA encoding a hum
33	36	3.0	1534	20 Z20856	Polynucleotide seq
C 34	36	3.0	1534	20 Z20856	Sequence encoding
35	36	3.0	4590	7 N60472	Sequence of the L.
C 36	35.8	3.0	1064	21 Z36918	DNA encoding Xenop
C 37	35.8	3.0	1348	21 Z65288	Human secreted pro
C 38	35.8	3.0	3705	20 X77691	Human transmembran
C 39	35.8	3.0	5852	12 O11710	Dictyostelium plas
C 40	35.8	3.0	8310	20 Z29911	CDNA encoding a SC
C 41	35.8	3.0	11459	20 X06876	Caenorhabditis ele
C 42	35.6	3.0	857	13 Q23373	CDNA coding for de
43	35.6	3.0	1172	15 Q86659	Protein allergen o
44	35.6	3.0	1502	20 V84577	Human secreted pro
45	35.4	3.0	251	19 X11360	Human diallelic po

ALIGNMENTS

RESULT 1	
ID V21209/c	V21209 standard; DNA: 1664976 BP.
XX V21209;	
AC V21209;	
XX 10-NOV-1990 (first entry)	
XX	
DE Methanococcus jannaschii circular chromosome.	
XX	
KW Methanococcus jannaschii; methanogenic archaeon; circular chromosome; genome; autotrophic; extrachromosomal element; Identification: ds.	
XX	
OS Methanococcus jannaschii.	
XX	
PN W09807830-A2.	
XX	
PD 26-FEB-1998.	
XX	
PF 22-AUG-1997; 97MO-US14900.	
XX	
PR 22-AUG-1996; 96US-0024428.	
XX	
PA (GENO-) INSR GENOMIC RES.	
PA (UNIT) UNIT ILLINOIS FOUND.	
PA (UIDO) UNIT JOHNS HOPKINS SCHOOL MEDICINE.	
PI Bult CJ, Smith HD, Venter JC, White OR, Woese CR;	
XX WPI: 1998-169145/15.	
DR	
XX	
PT Complete genome sequence of methano-genic archaeon, Methanococcus jannaschii - useful in identification of M. jannaschii genome	
PT fragment	
XX	

XX 18-JUN-1998.
 PD 11-DEC-1997; 97MO-US22787.
 XX 11-DEC-1996; 96US-0032757.
 PR (CHIR) CHIRON CORP.
 XX Escobedo J, Garcia P, Hu Q, Kothakota S, Williams LT;
 XX MPI; 1998-348453/30.
 DR P-PSDB; W63690.
 XX
 PT Secreted human polypeptides - having cytokine, cell proliferation or
 PT differentiation, activin or inhibin, tumour inhibition or
 XX anti-inflammatory activities
 XX
 PS Claim 6; Page 37; 78pp; English.
 XX
 CC This DNA encodes a human secreted protein. The specification provides
 CC secreted protein sequences (W63681 to W63699) encoded by the nucleic
 CC acid sequences shown in W43601 to W43619. The invention provides a
 CC method of identifying a secreted polypeptide which is modified by rough
 CC microsome. The secreted proteins can be used in assays to determine
 CC biological activities, such as cytokine, cell proliferation, or cellular
 CC differentiation activities, tissue growth or regeneration, or cellular
 CC inhibin activity, chemotactic or chemokinetic activity, haemostatic or
 CC thrombolytic activity, receptor/ligand activity, tumour inhibition, or
 CC anti-inflammatory activity. The proteins can also be used as biomarkers,
 CC or to identify tissues or cell types which express the proteins, or a stage-
 CC in disease-specific alteration in protein expression. They can be used
 CC in protein interaction assays, to identify ligands or binding proteins.
 CC Compounds which affect the biological activities of the secreted proteins,
 CC or their ability to interact with specific ligands can be identified
 CC using the proteins in screening assays. The proteins and antibodies that
 CC bind specifically to the protein can also be used to design diagnostic
 CC tests and therapeutic compositions for diseases which may be associated
 CC with altered expression of these proteins. Fusion proteins comprising
 CC e.g. signal sequences or transmembrane domains of the proteins can be
 CC used to target other protein domains to cellular membrane or they can
 CC be secreted extracellularly.
 XX
 SQ Sequence 1049 BP; 345 A; 214 C; 206 G; 284 T; 0 other;

Query Match
 Best Local Similarity 3.4%; Score 40; DB 19; Length 1049;
 Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1047 agctgataagatctgttaattcagttcagagagaataatgaattgtgcaactt 1106
 DB 891 agtgaatatttttaagctcttcttcaatgaagaacaaggcttactactt 950
 QY 1107 tcaattgacaaatcgaattgtatcttcttgcagtgatgaatttttaccataa 1166
 DB 951 tctctcattcattcgaatcaatcaaccattcaaaaaaataaaaaaataaaaaa 1010
 QY 1167 aaaaaaataaaaaa 1182
 DB 1011 aaaaaaataaaaaa 1026

RESULT 4
 200361
 ID 200361 standard; DNA; 2885 BP.
 AC 200361;
 XX 26-OCT-1999 (first entry)
 DE Nucleotide sequence of human hFATP6.
 XX

KW fatty acid transport protein; FATP; long chain fatty acid; LCFA; human;
 KW fatty acid, FATP biosynthesis; obesity; diabetes; heart disease; ss.
 XX Homo sapiens.
 OS MO9936537-A2.
 XX 22-JUL-1999.
 PD 14-JAN-1999; 99MO-US00182.
 PF 14-JAN-1999; 99US-0232201.
 PR 15-JAN-1998; 98US-0071374.
 PR 20-JUL-1998; 98US-0093491.
 PR 04-DEC-1998; 98US-0110941.
 PR 14-JAN-1999; 99US-0232195.
 PR 14-JAN-1999; 99US-0232197.
 PR 14-JAN-1999; 99US-0232200.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PA (WHD) WHITEHEAD INST BIOMEDICAL RES.
 XX
 PI Gimeno RE, Hirsch DJ, Lodish HF, Stahl A, Tartaglia LA;
 XX MPI; 1999-444398/37.
 DR P-PSDB; Y14951.
 XX
 PT Fatty acid transport proteins and related polynucleotides, useful
 PT for treating obesity, diabetes and heart disease
 XX
 PS Claim 2; Fig 54; 255pp; English.
 XX
 CC The invention provides a family of fatty acid transport proteins (FATPs)
 CC that mediate transport of long chain fatty acids (LCFAs) across cell
 CC membranes into cells. Human and murine FATP proteins and nucleic acids
 CC encoding the proteins are provided. The FATP proteins can be produced
 CC by standard recombinant methodology. Fatty acid uptake by cells can be
 CC modulated by regulating biosynthesis of FATP proteins especially FATP6.
 CC In particular, antisense oligonucleotides can be used to modulate FATP6.
 CC biosynthesis. Modulation of FATP6 is useful for inhibiting fatty acid
 CC uptake in cardiac muscle of humans. Agents can be directed to cardiac
 CC muscle or liver by administration of a complex of the agent and a FATP6
 CC binding moiety. DNA encoding FATP proteins can be used as a reference
 CC used in detecting variant alleles or homologues. Altering the LCFA uptake
 CC by administering an inhibitor or enhancer of FATP transport function in
 CC the small intestine can decrease or increase calories available as fats,
 CC and can decrease or increase circulating fatty acids. Blocking the
 CC function of FATP4 and also FATP2, is useful for treating obesity,
 CC diabetes and heart disease.
 XX
 SQ Sequence 2685 BP; 820 A; 557 C; 669 G; 839 T; 0 other;

Query Match
 Best Local Similarity 3.4%; Score 40; DB 20; Length 2885;
 Matches 100; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 981 agaaatgctctgtctgtcagtagagagcgcacatgctcgaatgtatgct 1040
 DB 2675 ataaacttagtgattattcttcttcttcttcttcttcttcttcttcttctt 2734
 QY 1041 acaatagctgataagatgttcaattcagttcagagagaataatgaattgtgac 1100
 DB 2735 ttccttaactaagaattcaataaataaataagctagcgcgttcgacaactac 2794
 QY 1101 aacttcatttgtaacaaatcgaattgtatcttcttgcagtgatgaatttttacc 1160
 DB 2795 aatgctccttccaataagaataattcgaataatttgataaagaattcaattctc 2854
 QY 1161 aatcaaaaaaataaaaaa 1180
 DB 2855 aaaaaaataaaaaaataaaaaa 2874

```

Db      1055 acccttaattaaagagatgatttlygcttataattcgatgagggaaacgcgatatttcg 1114
OY      1101 aacttttcattgatcatcaatgtatcttcttctgttcagtgatgaattttaccacaa 1160
Db      1115 atccctgaagtgatgaagagatttaacttataacttaattaaagctctgttctttaa 1174
OY      1161 aatcaaaaaaaaaaaaaaaa 1182
Db      1175 aaaaaaaaaaaaaaaaaaaa 1196

RESULF      6
X20517
ID      X20517 standard; DNA; 6761 BP.
XX
XX      X20517:
AC
XX
DT      05-MAY-1999 (first entry)
XX
DE      Polynucleotide sequence from the genome of Treponema pallidum.
KM      Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW      enzyme production; ds.
XX
XX      Treponema pallidum.
OS
XX      WO9859034-A2.
XX      PD
XX      30-DEC-1998.
XX
XX      23-JUN-1998; 98WO-US13041.
XX      PF
XX      24-JUN-1997; 97US-0050667.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      PI
XX      Fraser CM;
XX      WPI; 1999-081273/07.
DR
XX      New isolated Treponema pallidum nucleic acids - used to develop
PT      products for the detection, diagnosis, characterisation, prevention
PT      and therapy of T. pallidum infections, particularly syphilis
PS      Claim 1; Page 285-289; 1150pp; English.
XX
XX      X20500-21243 represent polynucleotide sequences from the genome of
XX      Treponema pallidum. The sequences can be used for detection,
XX      diagnosis, characterisation, prevention and therapy for T. pallidum
XX      infections, particularly syphilis. They can also be used for detecting
XX      diseases related to Borrelia infections in animals, and for the
XX      production of biosynthetic products such as enzymes.
CC
XX
XX      Sequence 6761 BP; 1693 A; 1585 C; 1749 G; 1723 T; 11 other:
SQ

Query Match      3.4%; Score 39.6; DB 20; Length 6761;
Best Local Similarity 53.5%; Pred. No. 0.97;
Matches 106; Conservative 0; Mismatches 89; Indels 3; Gaps 1
OY      483 cgaatttaactatagattgtccctgcaggtattgatgcccgggttataaccgcat 542
Db      2105 cgggtatcatattatcatatcatgatacaagcgctggtctgagaagaatgcgttaagctt 2164
OY      543 caccgctaacgaagccggtattgtattacaacaccgtatattactgcatitgagagacg 602
Db      2165 catctgcgcgaatgattcattgtgtgacacatcgccgaacctcagcgcaatcacacgatcg 2224
OY      603 atagagttacagcttctgttaattgta---tggaattcggatattaaatgattgta 659
Db      2225 atgagatgataaagaatcatcagcaactgaggttgatatactcggatatagaacttgaagatga 2284
OY      660 acagagttagaacatgatt 677

```

23-APR-2000 (first entry)
XX

XX *capricorn* annuum (Red pepper) ICC13 gene

iccid; red pepper; anthracnose disease; fruit crop; antifungal; antiviral; antibacterial; antinematodal; antiprotzoal; transgenic plant; pathogen resistance; phytopathogen; ss.

Capitulum annuum.

US6018038-A.

25-JAN-2000.

04-JAN-1999; 99US-0225244.

04-JAN-1999; 99US-0225244.

ANONIMO / NOROCCIO PETROCHEMICAL CO LTD.

on B, Kostenyuk I, Ko MK;

WPI; 2000-136708/12.
P-PSDB; Y67379.

P-PSDB; Y673379.

isolated cDNA clone from *Capsicum annuum* , encoding for an incomparable *Capsicum annuum* - *Colletotrichum gloeosporioides* interactions protein, for producing transgenic plants which exhibit enhanced resistance to phytopathogens -

CLAIM 1; Fig 1; 15pp; English.

An incomplete sequence represents a cDNA clone (designated *lcc13*), which encodes an incompatible *Capitulum annuum*-*Colletotrichum gloeosporioides* interactions (ICCI3) protein. Colletotrichum gloeosporioides is causal agent of anthracnose diseases on fruit crops, and shows incompatible interactions with ripe-red pepper fruit, and compatible interaction with unripe mature green fruit. The *lcc13* cDNA is unripe fruit *lcc13* expressed in the ripe fruit of the pepper, but not the anthracmated and antiprotazoal activity. The isolated cDNA clone is useful in an expression vector to produce a recombinant DNA expression system suitable for insection into cells to form a transgenic plant which exhibits enhanced resistance to phytopathogens; fungi, bacteria, viruses, nematodes, mycoplasma like organisms, parasitic higher plants, flagellate protozoa and insects.

sequence 1217 BP; 391 A; 221 C; 242 G; 363 T; 0 other;

Query Match	3.38;	Score 39.4;	DB 21;	Length 1217;
Best Local	Similarity	67.9%;		
Matches	55;	Conservative	0.51;	
		MicroMatch	0.51;	

acttaaaatattatattcttcaaaa
ccccccagcgaacgaaatcttactcaaa

1162 atCaaaaaa 1182

1185
1205

RESULT	8
Y2A385	

XX 10 standard; DNA; 2286 BP

AC X24385
xy

07-JUN-1999 (first entry)
D1
XX

XX gene promoter

stress tolerance; cold tolerance; drought tolerance; transgenic plant

XX XX Arabidopsis thaliana; ss

Arabis thaliana

FT	CDS	Location/Qualifiers
		2284...2296

```

/note="ERA1 gene start codon"

```

W09906580-A2.

11-FEB-1999

29-JUL-1998; 98WO-US15664

01-AUG-1997; 97US-0054474.

(PERF-) PERFORMANCE PLANTS INC

McCourt P, Ghassemian M, Culter S,

WPI; 1999-153807/13.

Disclosures: Fig 3; 66pp; English

The nucleotide sequence of the promoter region of the fatty acid transferase (FAT) ERAL gene of *Arabidopsis thaliana*, the coding region of the gene is provided in X24384, and codes for a 404-amino acid protein (see W97667) that is involved in the regulation of lateral branching, regulation of the response to abscisic acid (ABA) and regulation of senescence. The ERAL promoter is regulated in the guard cells of the plant and can affect water loss through the stomates. It can be used for the construction of gene products in guard cells. A claimed nucleic acid construct comprises a promoter and a nucleic acid encoding an inhibitor of a plant F₁. Such constructs can be used to produce plants having improved tolerance to drought, salt and cold stresses, reduced lateral branching, delayed senescence and increased numbers of flowers. Also claimed in a seed, plant part, cell or tissue culture or regenerated transgenic plant containing the claimed nucleic acid construct, the plant being a monocot or dicot, especially a Brassica sp. A plant having a mutation in the ERAL gene that results in loss of F₁ activity is also claimed.

Sequence 2205 BP; 791 A; 374 C; 434 G; 601 T; 6 other;

Query Match	3.38;	Score 39.2;	DB 20;	Length 2286;
Best Local Similarity	49.78;	Prod No 0;		

0; Mismatches 99; Indels 0; Gaps

1649 ++++++ 10

1046 aaagctgataaggaattgttttaattcaatcacttctt

Query Match	3.3%	Score 39;	DB 20;	Length 2328;
Best local Similarity	59.5%;	Pred. No. 0.87;		
Matches	66;	Conservative	0;	Mismatches 45;
			Indels	0;
			Gaps	0.

Query Match	3.3%	Score 38.8;	DB 19;	Length 1264;
Best Local Similarity	58.8%	Pred. No. 0.75;		
Matches	67;	Mismatches	47;	Gaps 0
		Conservative		

[illegible]

Tue Jan 9 08:34:57 2001

us-09-553-431-3.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2001, 09:20:42 ; Search time 81.49 seconds

(without alignments)
2337.606 Million cell updates/sec

Title: US-09-553-431-3
Perfect score: 1182
Sequence: 1 aagcttgatcgcacaccc.....tcaaaaaaaaaaaaaaaaaa 1182

Scoring table:

IDENTITY-NUC
Gapop 10.0, Gapext 1.0
Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/pdata/2/lna/5A_COMB.seq:*
2: /cgn2_6/pdata/2/lna/5B_COMB.seq:*
3: /cgn2_6/pdata/2/lna/6_COMB.seq:*
4: /cgn2_6/pdata/2/lna/PCUS_COMB.seq:*
5: /cgn2_6/pdata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.6	4.2	7218	1	US-08-232-463-14
2	45	3.8	7218	1	US-08-232-463-14
3	39.6	3.4	1223	3	US-09-154-874-4
4	39.4	3.3	1217	2	US-09-225-244-1
5	38.8	3.3	1264	3	US-08-758-621-13
6	38.8	3.3	1264	3	US-09-107-658-13
7	38.6	3.3	5852	1	US-07-867-106-2
8	37.2	3.1	2781	1	US-08-178-477B-42
9	36.4	3.1	1897	2	US-08-809-494A-1
10	36.4	3.1	1906	2	US-08-809-494A-1
11	35.6	3.0	857	3	US-08-460-040-1
12	35.6	3.0	1172	1	US-07-945-288-9
13	35.6	3.0	1172	1	US-08-462-831-9
14	35.6	3.0	1172	1	US-08-461-441-9
15	35.6	3.0	1172	1	US-08-461-441-9
16	35.6	3.0	1172	1	US-08-461-441-9
17	35.6	3.0	1172	1	US-08-461-441-9
18	35.6	3.0	1172	1	US-08-461-441-9
19	35.6	3.0	1172	1	US-08-461-441-9
20	35.6	3.0	1172	1	US-08-461-441-9
21	35.6	3.0	1172	1	US-08-461-441-9
22	35.6	3.0	1172	1	US-08-461-441-9
23	35.6	3.0	1172	1	US-08-461-441-9
24	35.6	3.0	1172	1	US-08-461-441-9
25	35.6	3.0	1172	1	US-08-461-441-9
26	35.6	3.0	1172	1	US-08-461-441-9
27	35.6	3.0	1172	1	US-08-461-441-9
28	35.6	3.0	1172	1	US-08-461-441-9

29	34.4	2.9	2894	3	US-08-884-324-14
30	34.2	2.9	2861	1	US-08-299-953-1
31	34.2	2.9	2861	1	US-08-299-953-1
32	34.2	2.9	2861	1	US-08-299-953-1
33	34.2	2.9	2861	1	US-08-299-953-1
34	34.2	2.9	2861	1	US-08-299-953-1
35	34.2	2.9	2861	1	US-08-299-953-1
36	34.2	2.9	2861	1	US-08-299-953-1
37	34.2	2.9	2861	1	US-08-299-953-1
38	34.2	2.9	2861	1	US-08-299-953-1
39	34.2	2.9	2861	1	US-08-299-953-1
40	34.2	2.9	2861	1	US-08-299-953-1
41	34.2	2.9	2861	1	US-08-299-953-1
42	34.2	2.9	2861	1	US-08-299-953-1
43	34.2	2.9	2861	1	US-08-299-953-1
44	34.2	2.9	2861	1	US-08-299-953-1
45	34.2	2.9	2861	1	US-08-299-953-1

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEITLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9p1-F15
US-08-232-463-14

Query Match

4.2% Score 49.6; DB 1; Length 7218;

Best Local Similarity 2.8%; Pred. No. 0.00019;
Matches 10; Conservative 207; Mismatches 141; Indels 0; Gaps 0;

556 agccgattagttacaacacctatattctcattgagagatgagatagttacagg 615
1449 AGAAGATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1390
616 ctgcttgaatgtatggaattagggatataaagtattgtaacagatgaaactga 675
1389 RR 1330
676 ttgataagggtgaagatagatgtcagttcttgcattgtaacagatgttgatgtc 735
1329 RR 1270
736 atgttgatgatacccgagatcgcagatgcagatgcagatgcagatgcagatgc 795
1269 RR 1210
796 tgtgtgaacagctccagctttagcagagatgcagatgcagatgcagatgcag 855
1209 RR 1150
856 ggtgagcaagatgacatgaagctgtgtatgtgagagaagaacttaagaagagga 913
1149 RR 1092

RESULT 2
US-08-232-463-14
Sequence 14, Application US/08232463

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
CLONE: pTZ9pT-Fls
US-08-232-463-14

Query Match 3.8%; Score 45; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 0.004;
Matches 15; Conservative 215; Mismatches 165; Indels 0; Gaps 0;

27 gatctctctctctcgcgcgagatgaccccgaggtctcaacagaaacctcactg 86
1064 GATYY 1123
87 tatgcatccacttcccaactcaatccclacacaaaccttaacccaacac 146
1124 YY 1183
147 caaacacctaccacaagccaccacaaatgcctcgtctcctcaatcacagcaacag 206
1184 YY 1243
207 agctgcgcgagagacacccacagatcgctgcaatcgacgcgagctgtgctgaacc 266
1244 YY 1303
267 legatctctctcgtctcgtaaaacgcgctcaatcacagctgtgtaagttcaacg 326
1304 YY 1363
327 gcgattgcagactgcacaaagccctagtcgtgataaagctgtgcaaatctgaattgc 386
1364 YY 1423
387 ttgtattcaaaactagtgtaaaatgccttta 421
1424 YYYYYYYYYYGACCAAAATCTTCTATCTCTTTA 1458

RESULT 3
US-09-154-874-4
Sequence 4, Application US/09154874

GENERAL INFORMATION:
APPLICANT: FADER, GARY MICHAEL
TITLE OF INVENTION: ISOFLAVONE BIOSYNTHETIC ENZYMES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154.874

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/931.668
FILING DATE: SEPTEMBER 17, 1997
ATTORNEY/AGENT INFORMATION:
NAME: MAJARIAN, WILLIAM R.
REGISTRATION NUMBER: 41,173
REFERENCE/DOCKET NUMBER: BB-1098-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4926
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 1223 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 55..1223
NAME/KEY: CDS
LOCATION: 55..990
US-09-154-874-4

Query Match
Best Local Similarity 3.4%; Score 39.6; DB 3; Length 1223;
Matches 78; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

OY 1041 acagtaagctgtaagagatgttttcaattcagatttcagagagaaattagaattgagc 1100
DB 1055 ACCCTTAATTTAAGGAGATGATTTGCTTATATTCGATGGGAACTGGATTTCGG 1114
OY 1101 aactttcattgatcaattcatttcttcttggttcagtgatgaatttactcaa 1160
DB 1115 ATCTTGAAATGTGAACGACTTTTAACTTATCAATTAATTAAGCTCTGCTTTTAAAAA 1174
OY 1161 aatcaaaaaaaaaaaaaa 1182
DB 1175 AAAAAAAAAAAAAAAAAA 1196

RESULT 4
US-09-225-244-1
; Sequence 1, Application US/09225244
; Patent No. 6018038
; GENERAL INFORMATION:
; APPLICANT: BOUNG-JUN OH
; APPLICANT: MOON KYUNG KO
; APPLICANT: IGOR KOSTENYUK
; TITLE OF INVENTION: Incompatible Plant and Pathogen Interaction Related Gene
; FILE REFERENCE: 1942/35
; CURRENT APPLICATION NUMBER: US/09/225,244
; EARLIER FILING DATE: 1999-01-04
; EARLIER APPLICATION NUMBER: No. 6018038 Yet Assigned
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Wordperfect 6.1 Windows
; SEQ ID NO 1
; LENGTH: 1217
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-225-244-1

Query Match
Best Local Similarity 3.3%; Score 39.4; DB 3; Length 1217;
Matches 55; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1102 actttcattgatcaattcatttcttcttggttcagtgatgaatttactcaa 1161
DB 1125 acttaaatatataatttcttcttggaagcaagtttaagatgagttatcttgaaa 1184
OY 1162 aatcaaaaaaaaaaaaaa 1182
DB 1185 aaaaaaaaaaaaaaaaaa 1205

RESULT 5
US-08-758-621-13
; Sequence 13, Application US/08758621
; Patent No. 5846821
; GENERAL INFORMATION:

APPLICANT: Guerino, Mary Lou, and Eide, David J.
TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,621
FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,578
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Silver, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: DCI-099CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1037
US-08-758-621-13

Query Match
Best Local Similarity 3.3%; Score 38.8; DB 2; Length 1264;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 1069 tcagtttcagagagaaattagaattgtagcaactttcatttgatcaattcattgla 1128
DB 1146 TTAGTTTGTATAGTCATGTGAAGCGTTTGAAGAAATTTCTTTATGACATTTGTAATTT 1205
OY 1129 ttcttgggttcagtgatgaatttactcaaaatcaaaaaaaaaaaaaa 1182
DB 1206 TATTTTATGATGCGATGTTTACTTCTTTAAAAAAAAAAAAAAAAAAA 1259

RESULT 6
US-09-107-858-13
; Sequence 13, Application US/09107858
; Patent No. 6162900
; GENERAL INFORMATION:
; APPLICANT: Guerino, Mary Lou et al.
; TITLE OF INVENTION: METAL-REGULATED TRANSPORTERS AND USES THEREFOR
; FILE REFERENCE: DCI-099CPDV
; CURRENT APPLICATION NUMBER: US/09/107,858
; EARLIER FILING DATE: 1998-06-30
; EARLIER APPLICATION NUMBER: 08/758,621
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1264
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:

3.38; Score 38.6; DB 1; Length 5852;

Qy 1093 attgtagcaacttttcatttgatcatlccaattglatittcttggttcagtgatgatattc 1152

Db 2691 ATTTACACACTCTATCTATCTTTTCATAAATAAATGTTAATGTCGCGTTGAGCTG 2750
 Oy 1153 ttaccacaatacaaaaaaaaaaaaaa 1182
 Db 2751 CAAATGAAAAAAAAAAAAAAAAAAAA 2780

RESULT 9

US-08-809-494A-1
 ; Sequence 1, Application US/08809494A
 ; Patent No. 5962260
 ; GENERAL INFORMATION:
 ; APPLICANT: Sawamura, Tatsuya
 ; APPLICANT: Masaki, Tomoo
 ; TITLE OF INVENTION: Modified low-Density Lipoprotein
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McAlay Fisher Nissen Goldberg & Kiel
 ; STREET: 261 Madison Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10016-2391
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/809,494A
 ; FILING DATE: 24-MAR-1997
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-321705
 ; FILING DATE: 30-NOV-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-214206
 ; FILING DATE: 31-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldberg, Jules E.
 ; REGISTRATION NUMBER: 24408
 ; REFERENCE/DOCKET NUMBER: JG-VY-4363PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 986-4090
 ; TELEFAX: 212 818-9479
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1897 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Bos taurus
 ; TISSUE TYPE: vascular endothelial cells
 ; IMMEDIATE SOURCE:
 ; LIBRARY: Bovine aortic endothelial cell cDNA
 ; CLONE: pBlox-1
 ; FEATURE:
 ; NAME/KEY: polyA_site
 ; LOCATION: 1880..1897
 ; FEATURE:
 ; NAME/KEY: misc_RNA
 ; LOCATION: 1859..1864
 ; OTHER INFORMATION: /function="polyA signal"
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: 1..34
 ; FEATURE:
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single

NAME/KEY: 3'UTR
 LOCATION: 848..1897
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 35..847
 US-08-809-494A-1

Query Match 3.1%; Score 36.4; Db 2; Length 1897;
 Best Local Similarity 47.1%; Pred. No. 0.58;
 Matches 112; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Oy 945 attggtgaatcgttgagttgggttggttggtgagagaatggtggttggttcacg 1004
 Db 1600 ATTGTTCAACTCATGAAAAAGTGTATTTGCTATTTAGATTTCTTATTTCTTTAA 1719
 Oy 1005 tagagctgcgtcgtgctacactgaaagtgtgtacagtaagctgataagattgtt 1064
 Db 1720 GCATGCTTTTATTTATGTTCTTTGTTGATGATTTGTTAGATTCATCATGATTTTCA 1779
 Oy 1065 taattcagtttcagagagaaatagaaatgtagtaacatttcatttcattcaat 1124
 Db 1780 TAATTATGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839
 Oy 1125 tgatctcttggttcagtgatgaaatttcacaaatcaaaaaaaaaaaaaa 1182
 Db 1840 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1897

RESULT 10

US-08-809-494A-3
 ; Sequence 3, Application US/08809494A
 ; Patent No. 5962260
 ; GENERAL INFORMATION:
 ; APPLICANT: Sawamura, Tatsuya
 ; APPLICANT: Masaki, Tomoo
 ; TITLE OF INVENTION: Modified Low-Density Lipoprotein
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McAlay Fisher Nissen Goldberg & Kiel
 ; STREET: 261 Madison Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10016-2391
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/809,494A
 ; FILING DATE: 24-MAR-1997
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-321705
 ; FILING DATE: 30-NOV-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-214206
 ; FILING DATE: 31-JUL-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldberg, Jules E.
 ; REGISTRATION NUMBER: 24408
 ; REFERENCE/DOCKET NUMBER: JG-VY-4363PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 986-4090
 ; TELEFAX: 212 818-9479
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1906 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/162,722
FILING DATE: 8-NOV-93
APPLICATION NUMBER: 07/458,642
FILING DATE: 13-FEB-90
APPLICATION NUMBER: PCT/AU88/00195
FILING DATE: 17-JUNE-88
APPLICATION NUMBER: PI 2523/87
FILING DATE: 18-JUNE-87
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMT-021CN2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..736
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 70..736

```

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 14
 US-08-461-809-9
 Sequence 9, Application US/08461809
 Patent No. 5770202
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
 NUMBER OF INVENTIONS: DERMATOPHAGOIDES
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 STATE STREET, SUITE 510
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII TEXT
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/461,809
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/945,288
 FILING DATE: 10 SEPTEMBER 1992
 APPLICATION NUMBER: US 580,655
 FILING DATE: 11 SEPTEMBER 1990
 APPLICATION NUMBER: US 458,642
 FILING DATE: 13 FEBRUARY 1990
 ATTORNEY/AGENT INFORMATION:
 NAME: MANDRAGOURAS, AMY E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1172 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single

LOCATION
US-08-461-441-9

Search completed: January 8, 2001, 09:20:56
Job time: 20531 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2001, 02:32:40 ; Search time 1361.86 seconds

(without alignments)
5032.326 Million cell updates/sec

Title: US-09-553-431-1
Sequence: 1 atggcgtctctcgtgagatgtt.....tctctctctctcttgagcgc 978

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_est4:*
5:	gb_est5:*
6:	gb_est6:*
7:	gb_est7:*
8:	gb_est8:*
9:	gb_est9:*
10:	gb_est10:*
11:	gb_est11:*
12:	gb_est12:*
13:	gb_est13:*
14:	gb_est14:*
15:	gb_est15:*
16:	gb_est16:*
17:	gb_est17:*
18:	gb_est18:*
19:	gb_est19:*
20:	gb_est20:*
21:	gb_est21:*
22:	gb_est22:*
23:	gb_est23:*
24:	gb_est24:*
25:	gb_est25:*
26:	gb_est26:*
27:	gb_est27:*
28:	gb_est28:*
29:	gb_est29:*
30:	gb_est30:*
31:	gb_est31:*
32:	gb_est32:*
33:	gb_est33:*
34:	gb_est34:*
35:	gb_est35:*
36:	gb_est36:*
37:	gb_est37:*
38:	gb_est38:*
39:	gb_est39:*
40:	gb_est40:*
41:	em_est41:*
42:	em_est42:*
43:	em_est43:*

44:	em_est44:*
45:	em_est45:*
46:	em_est46:*
47:	em_est47:*
48:	em_est48:*
49:	em_est49:*
50:	em_est50:*
51:	em_est51:*
52:	em_est52:*
53:	em_est53:*
54:	em_est54:*
55:	em_est55:*
56:	em_est56:*
57:	em_est57:*
58:	em_est58:*
59:	em_est59:*
60:	em_est60:*
61:	em_est61:*
62:	em_est62:*
63:	em_est63:*
64:	em_est64:*
65:	em_est65:*
66:	em_est66:*
67:	em_est67:*
68:	em_est68:*
69:	em_est69:*
70:	em_est70:*
71:	em_est71:*
72:	em_est72:*
73:	em_est73:*
74:	em_est74:*
75:	em_est75:*
76:	em_est76:*
77:	em_est77:*
78:	em_est78:*
79:	em_est79:*
80:	em_est80:*
81:	em_est81:*
82:	em_est82:*
83:	em_est83:*
84:	em_est84:*
85:	em_est85:*
86:	em_est86:*
87:	em_est87:*
88:	em_est88:*
89:	em_est89:*
90:	em_est90:*
91:	em_est91:*
92:	em_est92:*
93:	em_est93:*
94:	em_est94:*
95:	em_est95:*
96:	em_est96:*
97:	em_est97:*
98:	em_est98:*
99:	em_est99:*
100:	em_est100:*
101:	em_est101:*
102:	em_est102:*
103:	em_est103:*
104:	em_est104:*
105:	em_est105:*
106:	em_est106:*
107:	em_est107:*
108:	em_est108:*
109:	em_est109:*
110:	em_est110:*
111:	em_est111:*
112:	em_est112:*
113:	em_est113:*
114:	em_est114:*
115:	em_est115:*
116:	em_est116:*

117: em_estp16:*
 118: em_estp17:*
 119: em_estp18:*
 120: em_estp19:*
 121: em_estp20:*
 122: em_estp21:*
 123: em_estp22:*
 124: em_estp23:*
 125: em_estp24:*
 126: em_estp25:*
 127: em_estp26:*
 128: em_estp27:*
 129: em_estp28:*
 130: em_estp29:*
 131: em_estp30:*
 132: em_estp31:*
 133: em_estp32:*
 134: em_estp33:*
 135: em_estp34:*
 136: em_estp35:*
 137: em_estp36:*
 138: em_estp37:*
 139: em_estp38:*
 140: em_estp39:*
 141: em_estp40:*
 142: em_estp41:*
 143: em_estp42:*
 144: em_estp43:*
 145: em_estp44:*
 146: em_estp45:*
 147: em_estp46:*
 148: em_estp47:*
 149: em_estp48:*
 150: em_estp49:*
 151: em_estp50:*
 152: em_estp51:*
 153: em_estp52:*
 154: em_estp53:*
 155: em_estp54:*
 156: em_estp55:*
 157: em_estp56:*
 158: em_estp57:*
 159: em_estp58:*
 160: em_estp59:*
 161: em_estp60:*
 162: em_estp61:*
 163: em_estp62:*
 164: em_estp63:*
 165: em_estp64:*
 166: em_estp65:*
 167: em_estp66:*
 168: em_estp67:*
 169: em_estp68:*
 170: em_estp69:*
 171: em_estp70:*
 172: em_estp71:*
 173: em_estp72:*
 174: em_estp73:*
 175: em_estp74:*
 176: em_estp75:*
 177: em_estp76:*
 178: em_estp77:*
 179: em_estp78:*
 180: em_estp79:*
 181: em_estp80:*
 182: em_estp81:*
 183: em_estp82:*
 184: em_estp83:*
 185: em_estp84:*
 186: em_estp85:*
 187: em_estp86:*
 188: em_estp87:*
 189: em_estp88:*

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	408.4	41.8	457	140	C99848	C99848 C99848 YAC
2	408.2	41.7	710	93	AM668240	AM668240 GA_Ea001
3	341.4	34.9	873	94	AW729746	AW729746 GA_Ea002
4	214.4	21.9	557	107	BE359432	BE359432 DGL_40.A0
5	190.6	19.5	456	28	AL370067	AL370067 MRA35D01
6	181	18.5	276	174	B23006	B23006 F10H1F_1GF
7	169.2	17.3	588	107	BE359345	BE359345 DGL_40.A0
8	165	16.9	773	108	BE418872	BE418872 SCL083.D0
9	161.8	16.5	278	40	BE160136	BE160136 pB112 so
10	120.4	12.3	362	104	BE124676	BE124676 EST393711
11	100.8	10.3	104	165	AO958203	AO958203 LERAW015R
12	68.6	7.0	386	92	AW574469	AW574469 707092H09
13	51	5.2	463	170	A2303004	A2303004 GSSBR194
14	50	5.1	548	142	N37510	N37510 18737 Lambda
15	49.4	5.1	508	37	AV527004	AV527004 AV527004
16	47.2	4.8	1101	190	CNS010B7	AL098761 Drosophila
17	46.6	4.5	583	108	BE453489	BE453489 894071A05
18	44.4	4.5	880	190	CNS00KMD	AL078160 Drosophila
19	43.6	4.5	1101	190	CNS000D1	AL065414 Drosophila
20	43.6	4.4	1101	190	CNS01844	AL108862 Drosophila
21	42.2	4.3	588	105	BE194224	BE194224 HVSME008
22	40.4	4.1	959	190	CNS004XY	AL055406 Drosophila
23	39.6	4.0	939	190	CNS000CNG	AL064643 Drosophila
24	39.6	4.0	1101	190	CNS0182P	AL059400 Drosophila
25	39.4	4.0	507	108	BE462584	AL108811 Drosophila
26	39.4	4.0	551	93	AM626014	BE462584 EST324870
27	39.4	4.0	420	107	BE341746	AM626014 EST319921
28	39.2	4.0	1014	135	BE747742	BE341746 EST394577
29	38.8	4.0	1101	190	CNS0106X	BE747742 601578717
30	38.6	3.9	386	37	AV554133	AL098595 Drosophila
31	38.6	3.9	1101	190	CNS017XK	AV554133 AV554133
32	38.4	3.9	1101	190	CNS03931	AL108171 Drosophila
33	37.4	3.8	478	159	AO583931	AO543931 RPT-11-3
34	37.4	3.8	1021	190	CNS00DLS	AL072703 Drosophila
35	37.4	3.8	1221	190	CNS0169U	AL106476 Drosophila
36	37.2	3.8	1101	190	CNS0039G	AL065921 Drosophila
37	37	3.8	453	96	AW922683	AW922683 DGL_45-D0
38	37	3.8	561	94	AW746558	AW746558 WS1_54-A0
39	37	3.8	578	93	AM680908	AM680908 WS1_8_A11
40	37	3.8	581	97	AW924763	AW924763 WS1_72-G0
41	37	3.8	845	137	BE918164	BE918164 OVI_2-G02
42	37	3.8	845	190	CNS013KO	AL102978 Drosophila
43	37	3.8	1080	190	CNS00EPP	AL069494 Drosophila
44	37	3.8	1101	190	CNS00LOO	AL068607 Drosophila
45	37	3.8	1101	190	CNS0100X	AL098379 Drosophila

ALIGNMENTS

RESULT 1
 C99848
 LOCUS
 DEFINITION
 C99848 YAC clone C99848 region-specific CDNA Arabidopsis thaliana
 ACCESSION
 C99848
 VERSION
 C99848.1 GI:3719205
 KEYWORDS
 EST.
 SOURCE
 thale cress.

ORGANISM

Arabidopsis thaliana

REFERENCE

Arabidopsis thaliana
Brassicaceae; Arabidopsis.
1 (bases 1 to 457)
Ito, T.
Regional insertional mutagenesis of genes on Arabidopsis thaliana
chromosome V using Ac/Ds transposon in combination with a cDNA
scanning method
Unpublished (1998)
Contact: Takuya Ito
Laboratory of Plant Molecular Biology
The Institute of Physical and Chemical Research (RIKEN)
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Email: tallo@rc.riken.go.jp.
Location/Qualifiers
1. 457
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone_lib="YAC clone CTC811 region-specific cDNA"

JOURNAL

COMMENT

FEATURES

BASE COUNT 112 a 125 c 96 g 122 t
ORIGIN

Query Match 41.8%; Score 408.4; DB 140; Length 457;
Best Local Similarity 99.3%; Pred. No. 7.2e-117;
Matches 409; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atggcgctctgagatgcttcaagaacatcaatctctctccatcatcctc 60
DB 46 atggcgctctgagatgcttcaagaacatcaatctctctccatcatcctc 105
QY 61 tcaaaaaagacttaataatctcccaagatctgcaataccctagacagagatc 120
DB 106 tccaaaaagcctgtaataatctcccaagatctgcaataccctagacagagatc 165
QY 121 atagagtcgcttcaatcaatcaatcaatcaatcaatcaatcaatcaatca 180
DB 166 atagagtcgcttcaatcaatcaatcaatcaatcaatcaatcaatcaatca 225
QY 181 gtcgtatcaaccctcgaaagagcggtgttgaagaagacagacacacagagtc 240
DB 226 gtcgtatcaaccctcgaaagagcggtgttgaagaagacagacacacagagtc 285
QY 241 cctctctcgtcgttcaaggttctcagttgtcgcaatgagcgcagactgtctcgt 300
DB 286 cctctctcgtcgttcaaggttctcagttgtcgcaatgagcgcagactgtctcgt 345
QY 301 aacctgcatcctcctcctagagttagagatcgagatcaatcaatcaatcaatca 360
DB 346 aacctgcatcctcctcctagagttagagatcgagatcaatcaatcaatcaatca 405
QY 361 aacggaatctgtcgttcaatcaatcaatcaatcaatcaatcaatcaatca 412
DB 406 aacggaatctgtcgttcaatcaatcaatcaatcaatcaatcaatcaatca 457

RESULT 2

AM668240 710 bp mRNA EST 06-APR-2000
LOCUS GA_Ea0013E22 Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION
ACCESSION AM668240
VERSION AM668240.1 GI:7502620
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.

REFERENCE
AUTHORS
TITLE
An integrated analysis of the genetics, development, and evolution

JOURNAL

of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu

FEATURES

High quality sequence stop: 710.
Location/Qualifiers
1. 710
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="4400"
/db_xref="taxon:29729"
/clone_lib="Ea0013E22"
/issue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: PBK-CMV; Site_1: Ecom1; Site_2: XhoI"

BASE COUNT 185 a 157 c 213 g 155 t
ORIGIN

Query Match 41.7%; Score 408.2; DB 93; Length 710;
Best Local Similarity 76.6%; Pred. No. 9.6e-117;
Matches 500; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 118 ccaatcagtcgcttcaatcaatcaatcaatcaatcaatcaatcaatcaatca 177
DB 57 ccaatcagtcgcttcaatcaatcaatcaatcaatcaatcaatcaatcaatca 116
QY 178 atcgtcgtatcaaccctcgaaagagcggtgttgaagaagacagacacagagtc 237
DB 117 gtcgtcgtatcaaccctcgaaagagcggtgttgaagaagacagacacagagtc 176
QY 238 gtcgtcgtatcaaccctcgaaagagcggtgttgaagaagacagacacagagtc 297
DB 177 gtcgtcgtatcaaccctcgaaagagcggtgttgaagaagacagacacagagtc 236
QY 298 cgtaacctcagtcctcctcctcctcctcctcctcctcctcctcctcctcctc 357
DB 237 cgtaacctcagtcctcctcctcctcctcctcctcctcctcctcctcctcctc 296
QY 358 ataacgagagatctgtcgttcaatcaatcaatcaatcaatcaatcaatcaatca 417
DB 297 ataacgagagatctgtcgttcaatcaatcaatcaatcaatcaatcaatcaatca 356
QY 418 gaatctcgtatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 477
DB 357 gaatctcgtatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 416
QY 478 ttggaatgctgttgatgagtcgttgaaactgagcaggaagtacacagattcatc 537
DB 417 ttggaatgctgttgatgagtcgttgaaactgagcaggaagtacacagattcatc 476
QY 538 atcgaatgctgttgatgagtcgttgaaactgagcaggaagtacacagattcatc 597
DB 477 atcgaatgctgttgatgagtcgttgaaactgagcaggaagtacacagattcatc 536
QY 598 gcaatctcgttgatgagtcgttgaaactgagcaggaagtacacagattcatc 657
DB 537 gcaatctcgttgatgagtcgttgaaactgagcaggaagtacacagattcatc 596
QY 658 ttgttagaagcagatgagatcagagatcaataagagatgagatgagatgagatg 717
DB 597 ttgttagaagcagatgagatcagagatcaataagagatgagatgagatgagatg 656
QY 718 atgattaaagagagatgagatgagatgagatgagatgagatgagatgagatg 770
DB 657 atgattaaagagagatgagatgagatgagatgagatgagatgagatgagatg 709

```

RESULT 3
LOCUS   AM729746
DEFINITION  GA_Ea0025013 Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum CDNA clone GA_Ea0025013, mRNA sequence.
ACCESSION  AM729746
VERSION    AM729746.1
KEYWORDS   GI:7627348
SOURCE     Gossypium arboreum.
            Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Malvales; Malvaceae; Gossypium.
REFERENCE  1 (bases 1 to 873)
AUTHORS   Leslie, A., Frisch, D., Yu, Y., Wood, T. C., Wing, R. A. and Wilkins, T. A.
TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
JOURNAL   Unpublished (2000)
COMMENT   Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            High quality sequence stop: 873.
            Location/Qualifiers
                1..873
                /organism="Gossypium arboreum"
                /strain="AKA"
                /cultivar="8400"
                /db_xref="taxon:29729"
                /clone="GA_Ea0025013"
                /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                /tissue_type="Fibers isolated from bolls harvested 7-10
                dpa"
                /lab_host="E. coli"
                /note="vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT  216 a 186 c 268 g 203 t
ORIGIN
Query Match          34.9%; Score 341.4; DB 94; Length 873;
Best Local Similarity 72.9%; Pred. No. 8.6e-96;
Matches 466; Conservative 0; Mismatches 171; Indels 2; Gaps 2;

OY 118 ccaatacgaatcggtcttcaattaatcgaacacggaactcgcgagagaacgcgcgt 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 56 CCCATTCTTCTGCTGCTGATGAAATAAAGCTGAGCTGGCGGGGAGACTCCAGC 115

OY 178 atcgctgtatataccctccggaagagcggtgttg--aaagcgacacacccgcaaatgt 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 116 GCGTCGTATATACATCCGGAAGAGCGGGGTCGTAAACCAACCAACACCGCAATGT 175

OY 237 cggctctctctgcgctgaacggttcctcagttgtctgcacatgacgcgcgacctgtgtc 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 GGGCTCTTACTTACCTGCTGCTGGGCTTCTCCGTAAGTCCCATGACGCCGACGCTTGGCCT 235

OY 297 ccgtaacct-cgaatctctctcctagaggttagaagatcgatcaatctacatgtcgtagg 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 ACATAACCATATGACCTCTCTTATGCTTGAGAAATCCGCTGAACCTATATCTCTCTGAGG 295

OY 356 tataaagagagatgtgtcgtcgatcaactctgctgacgtgataagcgttgagcaatt 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 TACTCATATGCGGATGCGGACTGACCATCATTTGGTAAGATGAAGATGATGTCGAACT 355

OY 416 tcgaatgtatgtatatacacaactagatcgaacttcgattgagatgttggtgtaag 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 TATTAATTTGCTTTCATTTCACAGCCAGATCGAAGATACCATTTGATTCAGTGCGGAAG 415

OY 476 catgtgaatgagcttgatgagcttgaaactagaccggaaggttcacggaattcatca 535

```

```

DB 416 CGTTGGTTTGGCTAGTGGACCCACTGAATGCGCGGAGAGAGGCTACCGGATTTTCATAC 475
OY 536 tcatcagatgctcctgcaggaatcgatcggaattcatcaaccgcaattactccgcgaatg 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 476 TATATATTTGATGACGGCGGGGATTAACGCGAGGTTTCATTAACGCCATTAACGCTAACG 535
OY 596 aagacatctcgtgtacacaccccgatataacacgcttaaggagatgctgataaggttacg 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 536 AGCGGCTTTTGGTGCACGACCGCTGACATTAACAGCTTAATGATCTCATAGAGTACAC 595
OY 656 gttcttgaagatcgatgaaatcagagatacaaaagatgattgtgacacagatgagactg 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 596 GCGCTTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 655
OY 716 atatgattaaagagagagatgattgtatgtcagtgtagatg 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 656 CTCTGATTAGAGGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 694

RESULT 4
LOCUS   BE359432
DEFINITION  Dcl_40_A03.91_A002 Dark Grown 1 (Dg1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION  BE359432
VERSION    BE359432.1
KEYWORDS   GI:9300989
SOURCE     EST.
            sorghum.
ORGANISM   Sorghum bicolor
            Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.
REFERENCE  1 (bases 1 to 557)
AUTHORS   Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt
            , L.H.
TITLE     An EST database from Sorghum: dark-grown seedlings
JOURNAL   Unpublished (2000)
COMMENT   Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmp@prattuga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTmix
            High quality sequence start: 6
            High quality sequence stop: 531
            POLYA-NO.
FEATURES
    source
        1..557
        /organism="Sorghum bicolor"
        /db_xref="taxon:4558"
        /clone_lib="Dark Grown 1 (Dg1)"
        /note="Organ: 5-day-old dark-grown seedlings. Vector:
        lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
        made from poly-A RNA in the cloning vector lambda Zap II.
        Clones to be sequenced were prepared by mass excision."
BASE COUNT  116 a 127 c 174 g 140 t
ORIGIN
Query Match          21.9%; Score 214.4; DB 107; Length 557;
Best Local Similarity 69.7%; Pred. No. 4.4e-56;
Matches 290; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

OY 523 ccgattatcatcatcatgctcctgcaggaatcgatgacggaattcataaccgcaatt 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 CCCGCTTCATCATCTCATGACATGCCGCCGACAGGTGTGATGCCGGTTGTGCTACTGCATT 60

OY 583 actcggcgaatgaagcagctctctgtaacaactccggaatataacagcgttaaggaatgct 642

```

Db	61	GCACCTGCAGAAAGACGACACTGCTCTCTTCCACCCCTGCATCTTACGGCTCTCCGGTATGCT	120
Qy	643	gataaggttcaagggttltgtctaaagatcgaatcgaatcagaatataaagaatgatcttgaac	702
Db	121	GACCGTGTCCGAGGACTGTTGTGAGTGCATGCGATCAAAAGTATCCAAAGTATTATTGTAAC	180
Qy	703	agaagtgaagaccgcatgatgatataaagaagaagatgattgtcaatgttaagatgattgacgaag	762
Db	181	CGAGTGCAGCCACGACCTGCTGTAAGGGGAGAGACATGATGTGAGACTTGTATGTCCAAAGAA	240
Qy	763	atgttggagatgtcatcttctgtgtgttaattctctgaagaattcgaattatctgaacgaacg	822
Db	241	ATGCTTGGCGTTCGCCCTTGCTTGGTGTGCTCTCAGATGACGGAACTAAATCCCGAGTACA	300
Qy	823	aatcgaagggttccgcttctgtctgaatlaaagccctctcagctcttgcggagatctgagcttgg	882
Db	301	AATAGGCGGTACCATTTGGTGTCTCAACAGACCCGCCACAGCGCTCCGGCGCTTCCTGGAC	360
Qy	883	caggcggcttggagaaactcgtttgagcaagaatagatlaaagcgttctatgttggtaaga	938
Db	361	CAGGCTTACTTGGCGATTGCTGGAAAGAGATGATGATGACAGCATCTATGCTTTAAGA	416

RESULT	5
LOCUS	AL370067
DEFINITION	AL370067 456 bp mRNA
ACCESSION	M8BA35D01F1 M8BA Medicago truncatula EST
VERSION	AL370067 sequence.
KEYWORDS	AL370067.1 GI:9669820
SOURCE	EST.
ORGANISM	barrel medic. Medicago truncatula

FEATURES	source
Location/Qualifiers	1. .456

BASE COUNT	110 a	68 c	141 g	137 t	ORIGIN
Query Match	19.5%	Score 190.6;	DB 28;	Length 456;	
Best Local Similarity	81.8%;	Pred. No. 1.2e-48;			
Matches 220;	Conservative 0;	Mismatches 49;	Indels 0;	Gaps 0;	
Qy	517	ggttaaccggatattcattcattcattcgaattgctccagagaatcgaatccggaattcataacc	576		
Db	188	GGCGTCGCGCACTTATCTCGATCGATTCGCCGCGGATTCGATTCGGGTTTAAACG	247		
Qy	577	gacattactcccgagataaagcagttctcgttaacaactccgagatalaacaacgcttaagg	636		
Db	248	GGCATTAACACCCGGCTNATTAAGCGGTTTGAATACACCCCGGATATAATCGCGTTGAGG	307		
Qy	637	gattctgataagggttaacgggttcgtttaagatgcgaatggaatcaagaatataaagaatgatt	696		
Db	308	GATGCGGATGAGAGTAGCGGAGTTGTTGGAGTGTGTGCGATTTAGCGATATTAAAGATCATTA	367		
Qy	697	gtgaacagagatgaagacatgatatgataaagaagaagataatgaatcagttcttgatggtg	756		
Db	368	GTCGATTAGGGTTTAGACGCGATATATGATTTAAAGTGAGATATGATGTCGGTTTTGGATGTG	427		
Qy	757	caggagatgttgggaattgcatctgctgtg	785		
Db	428	CAAGGATGTGGGTTTGCCCTTCCTTGG	456		

RESULT	6
B23006/C	
LOCUS	B23006 276 bp DNA
DEFINITION	F16H1TR IGF Arabidopsis thaliana genomic clone F16H1, DNA sequence
ACCESSION	B23006
VERSION	B23006.1 GI:2507810
KEYWORDS	GSS.
SOURCE	Thale cress.
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
	D Brassicales; Brassicaceae; Arabidopsi.
	1 (bases 1 to 276)
REFERENCE	Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Coffey,E., Golden,K.
AUTHORS	, Johnson,K., Adams,M.D. and Venter,J.C.
TITLE	A BAC End Sequence Database for Identifying Minimal Overlaps in
JOURNAL	Arabidopsis Genomic Sequencing
COMMENT	Unpublished (1997)
	Other GSSs: F16H1R

BASE COUNT	ORIGIN
77 a	76 c
53 g	70 t

Query match

18.5%; Score 181; DB 174; Length 276;

[illegible]

Query Match	16.9%;	Score 165;	DB 108;	Length 773;
Best Local Similarity	60.1%;	Pred. No. 1.5e-40;		

Matches	326:	Conservative	0:	Mismatches	211:	Indels	5:	Gaps	3:
OY	151	ccggacacccgcgcggagaacccgcggtatctgctgcgtatcactccgcggaaagcggtgt	210						
Db	99	CCGGAGCTGTGGGCGCCCGACCCGCGGGGTGTGGTGTGACCTCGGGGCAAGGGCGCGTG	158						
OY	211	ggaagagcagacaaccacgcgcaaatgctgcgtctctctcgcgtcgtgtacggttctcagt	270						
Db	159	GGGAAGACACACACACCGCCCAACCTGGCCCGCCCTGCCGTGGCGCGCGGCGCTCCCGC	218						
OY	271	gtcgacattgcgcgcgcacattgctctgcgttaaccctgcgtctctctctggttaagaat	330						
Db	219	GTGGCCCTGTGGAGCCCGACCGGGCCTCCGGAACTCGACCTGCTCTGCGGCTGGAGAAC	278						
OY	331	cgaagcaaltcaactctgctgcgtggttataacgagaaattgctcctcgtatcaagctctg	390						
Db	279	CGGGTGAACCTGTACCCGCGCGACCTGCTGCCGGCGGCACTGCCCGCTCGACACGCCCTC	338						
OY	391	gtacgtgaataagcgtctgctgaatttcgtaattgcatatatactcaaacctatgacgaa	450						
Db	339	GTCCGCCACACCGCTCCCTCCGGGGCCTCCACCTGCTCTGCTCTCCAAAGCCCGCTCCAG	398						
OY	451	cttcgcgatggaatttgtgtgtaagcattggaatggtcttgatgctgttgaanaactaga	510						
Db	399	CTGCCCTTCGCTTCGGCTCCAGAGCCCTTCACCTGGTGGCGGACGCCCTCCGCGCGCG	458						
OY	511	ccggaaggttcacgcggaattcatcatcatgcatgctctcgcaggaatcgtatccggaatc	570						
Db	459	GCCGA---CCCGCCCGCCTTCATNCTCATGCACCTCCCTGCAGGCTTGATCTTGCGTTT	515						
OY	571	ataacgcacatlaactcgcggcggaatgaagcagtlctcgttaacaactcgcgatalaaccgc	630						
Db	516	GTCACAGCCATTCGCCCTTCACAGAAC-ACCAGGCTGTGGTACCACCTCGACATCCAC	573						
OY	631	ttaaggaatgcgtcgaatgaaggttaacggttctgttagaatcgtatcgatcagatcagaataa	690						
Db	574	CTCGTGATGCAGACGGCGTCCGGGGCTGTGGATGTGACGGCATCAAGACATCAAG	633						
OY	691	at 692							
Db	634	at 635							
RESULT	9								
LOCUS	AW160136	278 bp	mRNA	EST	01-MAY-2000				
DEFINITION	polh12 soybean, century cDNA library Glycine max cDNA similar to								
ACCESSION	AW160136								
VERSION	AW160136.1	GI:7673729							
KEYWORDS	EST.								
SOURCE	soybean.								
ORGANISM	Glycine max								
REFERENCE	Eukaryote: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;								
AUTHORS	Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids I;								
	Fabales; Fabaceae; Papilionoideae; Glycine.								
	1 (bases 1 to 278)								
	Matthews,B.F., Devine,T.E., Weisemann,J., Beard,H.B., Lewers,K.S.,								
	MacDonald,M.H., Park,Y.B., Maitl,R., Lin,J.J., Kuo,J., Pedroni,M.J.A.,								
	Cregan,P.B. and Saunders,J.A.								
	A soybean genetic map incorporating cDNA and EST markers and								
	anchored with SSR markers								
TITLE	Unpublished (2000)								
JOURNAL	Contact: Beard HB								
COMMENT	Soybean and Alfalfa Research Lab								
	USDA ARS PSI								
	10300 Baltimore Ave, Bldg 006 Rm115, Beltsville, MD 20705, USA								
	Tel: 3015045376								
	Fax: 3015045728								
	Email: hbeard@ars.arsusda.gov								
	Insert length: 500 Std Error: 0.00								
	High quality sequence stop: 278								
	POLYA-No.								

FEATURES		source		Location/Qualifiers	
		1. 278			
		/organism="Glycine max"			
		/cultivar="Century"			
		/db_xref="taxon:3847"			
		/clone_lib="soybean, century cDNA library"			
		/note="Vector: pBluescript SK-"			
BASE COUNT	64 a	58 c	88 g	67 t	1 others
ORIGIN					
Query Match	16.5%	Score 161.8;	DB 40;	Length 278;	
Best Local Similarity	76.0%	Pred. No. 11e-39;			
Matches 199;	Conservative	0;	Mismatches 63;	Indels	0; Gaps 0;
QY	703	agaagtcgagactgatatattaaagagagagatacgaatcagttgtatgattgcagag	762		
Db	17	AGGGTTCGGACCGACGATCATCAAGGGAGGACATGATGTGGTGTTGGAGCTGCACAG	76		
QY	763	atgtcggagatgcatcttgcttgggtgaattctctgaagatctgaagttatccgaagcag	822		
Db	77	ATGTATGGGCTTGGCCCTTCTCTCGGGGTTATTCCTGAGGATAGAGAGTTATTAAGAAGC	136		
QY	823	aatcgaaggttcgcgcctgtctcgaataaagcctccctacagcttcgaggaattgcgttga	882		
Db	137	AATAGAGGTTTCCCTCTGTGTGCTCAACAAGCCTCCACATTTTGGCGGATTTGGCGTTG	196		
QY	883	cagcgtgccttggagactcgttgagcaagaatgatatgaagcgtgttatgttgaggaaga	942		
Db	197	CNAGCGCGGTGAGAGCTGTGTGAGCAAGATGACATGACGAGCGCGNGTGTGTGAAACNA	256		
QY	943	cctaagaacgtgcgctctctct 964			
Db	257	CCCAACGTCGGCTTTTCTCT 278			
RESULT 10					
BE124676					
LOCUS	BE124676	362 bp	mRNA	EST	07-SEP-2000
DEFINITION	EST393711	GVN	Medicago truncatula	cDNA clone	pgVN-67C12, mRNA
ACCESSION	BE124676				
VERSION	BE124676.1	GI:8529233			
KEYWORDS	EST.				
SOURCE	bairrel medic.				
ORGANISM	Medicago truncatula				
REFERENCE	Eukaryotes; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoidae; Medicago.				
AUTHORS	1 (bases 1 to 362) Fedorova,M., Pearson,B.L., Samac,D.A., Vance,C.P., Gantt,J.S., Peng,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.				
TITLE	ESTs from one month old nitrogen-fixing root nodules of Medicago truncatula				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Carroll P. Vance Department of Agronomy and Plant Genetics University of Minnesota 411 Borling Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA Tel: 612-625-7219 Fax: 651-649-5058 Email: vance004@maroon.tc.umn.edu University of Minnesota name:M261544e TIGR sequence name:MTCCB187K More information is available at: http://chrysis.tamu.edu/medicago Seq primer: SKmod (CTA gaa cta gctg gat cc).				
FEATURES	Location/Qualifiers				
SOURCE	1. 362 /organism="Medicago truncatula" /cultivar="genotype A17" /db_xref="taxon:3880"				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2001, 03:36:26 ; Search time 2611.99 Seconds

(Without alignments)
1916.225 Million cell updates/sec

Title: US-09-553-431-1
Sequence: 1 atggcgcctctgagatcgtt.....tcttcttcttcttgcgcgc 978

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_om:*
6: gb_ph:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*
42: gb_ph:*
43: gb_ph:*

44: em_hlg8:*
45: em_hlg9:*
46: em_hlg10:*
47: em_hlg11:*
48: em_hlg12:*
49: em_hlg13:*
50: em_hlg14:*
51: em_hlg15:*
52: em_hlg16:*
53: em_hlg17:*
54: em_hlg18:*
55: em_hlg19:*
56: em_hlg20:*
57: em_hlg21:*
58: em_hlg22:*
59: em_hlg23:*
60: em_hlg24:*
61: em_hlg25:*
62: em_hlg26:*
63: em_hlg27:*
64: em_hlg28:*
65: em_hlg29:*
66: em_hlg30:*
67: em_hlg31:*
68: em_hlg32:*
69: em_hlg33:*
70: em_hlg34:*
71: em_hlg35:*
72: em_hlg36:*
73: em_hlg37:*
74: em_hlg38:*
75: em_hlg39:*
76: em_hlg40:*
77: em_hlg41:*
78: em_hlg42:*
79: em_hlg43:*
80: em_hlg44:*
81: em_hlg45:*
82: em_hlg46:*
83: em_hlg47:*
84: em_hlg48:*
85: em_hlg49:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	978	100.0	981	AB030278	AB030278 Arabidops
2	978	100.0	65958	AB009056	AB009056 Arabidops
3	452.6	46.3	1164	AF251019	AF251019 Tagetes e
4	247.4	25.3	118360	AF166114	AF166114 Mesostigm
5	243.6	23.0	150613	AB001684	AB001684 Chlorella
6	224.6	23.0	200799	AF137379	AF137379 Nephrosel
7	224.6	22.9	121524	AF041468	AF041468 Gullardi
8	224.2	22.0	22085	PW1245645	PW1245645 Proteobac
9	214.8	19.9	135638	33 SYCSLRG	33 SYCSLRG
10	194.2	19.6	194509	7 AP001129	7 AP001129 Oryza sat
11	192	13.7	673	AF149810	AF149810 Oryza sat
12	133.6	13.3	14103	1 AE000772	1 AE000772 Aquifex a
13	130.4	13.2	1900	2 BSMNCD	2 BSMNCD
14	128.8	13.2	4389	2 BACREMIN	2 BACREMIN
15	128.8	13.2	5641	2 BACDIREG	2 BACDIREG
16	128.8	13.2	218410	2 BSUB0015	2 BSUB0015
17	128.8	12.5	294250	2 AP001517	2 AP001517 Bacillus su
18	122.4	11.2	16486	1 AE001931	1 AE001931 Deinococc
19	109.4	10.4	8039	1 AE001824	1 AE001824 Thermotog
20	101.4	9.0	12165	1 AE003965	1 AE003965 Xylella f
21	87.6				

C	22	81.4	8.3	21607	1	AE001468	AE001468 Helicobac
C	23	79.8	8.2	13818	1	AE004747	AE004747 Pseudomon
C	24	79	8.1	2400	2	ECOM1B	J03153 E. coli mini
C	25	79	8.1	11998	1	AE000216	AE000216 Escherich
C	26	79	8.1	17045	2	D90751	D90751 Escherichia
C	27	79	8.1	17447	2	D90752	D90752 Escherichia
C	28	78.8	8.1	340806	2	NMA12491	NMA12491 Neisseria
C	29	76.2	7.7	11184	1	AE004271	AE004271 Vibrio ch
C	30	75	7.7	12907	1	AE000551	AE000551 Helicobac
C	31	74	7.6	9966	1	AE002374	AE002374 Neisseria
C	32	67.4	6.9	994	77	CNS01KX	AL148106 Anopheles
C	33	65	6.6	293181	2	AP001119	AP001119 Buchnera
C	34	57.2	5.8	12730	1	AE005101	AE005101 Halobacte
C	35	57	5.8	110071	1	AE000937	AE000937 Methanoba
C	36	53.6	5.5	11490	1	AE005032	AE005032 Halobacte
C	37	50.6	5.2	10699	29	U67504	U67504 Methanococc
C	38	49.4	5.1	45013	7	F22M8	AC020622 Sequence
C	39	46.6	4.8	7218	81	166494	Sequence 14
C	40	46.6	4.8	319000	2	AP000006	AP000006 Pyrococcu
C	41	46.2	4.7	9722	1	AE002534	AE002534 Neisseria
C	42	44.8	4.6	11813	1	AE001172	AE001172 Borrelia
C	43	44.8	4.6	13494	1	AE001056	AE001056 Archaeogl
C	44	44.6	4.6	293250	2	CNSPAX02	AJ248284 Pyrococcu
C	45	42.8	4.4	12637	1	AE005148	AE005148 Halobacte

ALIGNMENTS

RESULT 1
LOCUS AB030278 981 bp mRNA PLN 11-FEB-2000
DEFINITION Arabidopsis thaliana mRNA for Mind, complete cds.
ACCESSION AB030278.1 GI:6759276
VERSION
KEYWORDS Mind.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
Arabidopsis.

REFERENCE
AUTHORS Kim, M., Fujiwara, M., Kanamaru, K., Tanaka, K. and Takahashi, H.
TITLE Arabidopsis thaliana mind homolog involved in plastid division
JOURNAL Unpublished (1999)
REFERENCE 2 (bases 1 to 981)
AUTHORS Kim, M., Fujiwara, M., Kanamaru, K., Tanaka, K. and Takahashi, H.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-1999) to the DDBJ/EMBL/GenBank databases. Hideo
Takahashi, Institute of Molecular and Cellular Biosciences, The
University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-0032, Japan
(E-mail: htakahashi@mol.f.u-tokyo.ac.jp/genetics/mgb.html,
URL: http://imcbs.fam.u-tokyo.ac.jp/genetics/mgb.html,
Tel: 81-3-5841-7825, Fax: 81-3-5841-8476)

FEATURES
source
1..981
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="v"
1..981
/gene="mind"
1..981
/gene="mind"
/function="mind homolog involved in plastid division"
/codon_start=1
/product="Mind"
/protein_id="BA090261.1"
/db_xref="GI:6759276"
/translation="MASLRIFSTNHOSLLPSSLSOKTILSSPRFVNPSRSPISRV
LOPNRDELAGEPRIVITSCGKGVGKTTTANVGLSLARVGSVATDADGLNL
DILIGENRYNTCEVINGDCRLDQALVDRKWSNELCLISPKSPKIPMGEGKAL
EMLVDAKTRPESPDFIITDCPADIGAFITAITPANEAVLVTTPIALRDADRYT

Query Match 100.0%; Score 978; DB 6; Length 981;
Best Local Similarity 100.0%; Pred. No. 3.8e-271;
Matches 978; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
BASE COUNT 253 a 205 c 249 g 274 t
GILECDGIRIDIKMIVNRYRTDMIGEDMMSVLDVQEMILGLIAGVIPEDSEVIRSTNR
GFFLVLPKPTPLAGLAFQDAAMRIYEDQSMKAVMYEPEPKRGFFSFFG"

QY	1	atgcgcgtctgagatgtgtctcaacgaatcatcatcttcttcctcatcatctc	60
DB	1	ATGGCGCTCTGAGATTGTTCTCAACGAATCATTAATCTTCTCTCCATCATCTCTC	60
QY	61	tcacaaagactctcaatctctcaacgaatctgcaataaccctagcaagagatcca	120
DB	61	TCACAAAAGACTCTCAATCTCTCAACGAATTCGTAATACCTAGAGAGAGTCCA	120
QY	121	atacagtcgtctctcaatctcaatcgaaccggaactcgcggaagaacgcgcatc	180
DB	121	ATACAGTCTCTCTCAATCTCAATCGCAACCGGAGTCCGGAGAAACCGCGTATC	180
QY	181	gtcgtatcaccctccggaagaagcggtgttggaagaagacacacacgcaatgctg	240
DB	181	GTCGTATCACTCCGGAAGAAGCGGTGTGGAAGAAGACACACCAACCAATGTCGT	240
QY	241	ctctctcgtcgtctgctggtttctcagttgctgcgcatgtagcgccgaccttgctc	300
DB	241	CTCTCTCGTCTGCTGTACCGTTTCAGTTTCGTCATTCAGCGACCTTGCTCGCT	300
QY	301	aaactcgatctctctctagagtagagaaatcgatcaatctgclcgagagttata	360
DB	301	AACCTCGATCTCTCTAGAGTAGAGAAATCGATCAATCTGCTCGAGTTATA	360
QY	361	aacgagatgtgctcgcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga	420
DB	361	AACGAGATGTGCTCGCATCAAGCTGTGATGATGATGATGATGATGATGATG	420
QY	421	ttgctatgataatctcaacgaatcgaatcgaatcgaatcgaatcgaatcgaatc	480
DB	421	TTCGTAATGATATCTTAACCATCAAGATCAATCTCGATGATGATGATGATG	480
QY	481	gaatgctgtgtagtgcgttgaaacacgaacgaacgaacgaacgaacgaacgaac	540
DB	481	GAATGCTGTGAGTGCCTTGAACACGATGACCGGAGGATTCACCGATTCATC	540
QY	541	gattgctcgtcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga	600
DB	541	GATTGCTCTCGAATGATGATGATGATGATGATGATGATGATGATGATGATG	600
QY	601	gttcgtgaac	660
DB	601	GTTCTGTGAAC	660
QY	661	ttagaagatgagatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga	720
DB	661	TTAGAATGAGATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGA	720
QY	721	attaaagagagatagatgttcagtgtagatgttcagagatgttcagatgttc	780
DB	721	ATTAAAGAGAGATAGATGTGATGTGATGTGATGTGATGTGATGTGATGTG	780
QY	781	cttggtgtaattctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga	840
DB	781	CTTGGTGTAATCTCTGAATCTCTGAATCTCTGAATCTCTGAATCTCTGA	840
QY	841	gttcgaataagcctcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatc	900
DB	841	GTTCTGAATAGCCTCTACGCTGCGGATGCGGATGCGGATGCGGATGCG	900
QY	901	gttgagcaagatagatgagagagctgtagttagttagttagttagttagttag	960
DB	901	GTTGAGCAAGATAGATGAGAGGCTGTATGTGTGAGGAGAACTAAGAAACG	960

QY 961 ttctctctcttgccgagc 978
 Db 961 TTCTCTTCTTTTGCCGC 978

RESULT 2
 AB009056
 LOCUS AB009056 65958 bp DNA 09-AUG-2000
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MZF18.
 ACCESSION AB009056
 VERSION AB009056.2 GI:97582219
 KEYWORDS
 SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_11b:Mitsui p1
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Sato,S., Kaneko,T., Kotani,H., Nakamura,Y., Asamizu,E., Miyajima,N.
 and Tabata,S.

REFERENCE
 AUTHORS
 TITLE Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 Sequence features of the regions of 1,456,315 bp covered by
 nineteen physically assigned P1 and YAC clones
 DNA Res. 5 (1), 41-54 (1998)
 JOURNAL 98290546
 MEDLINE 2 (bases 1 to 65958)
 REFERENCE Nakamura,Y.
 AUTHORS Direct Submission
 JOURNAL Submitted (27-NOV-1997) to the DDBJ/EMBL/GenBank databases.
 Yaskazu Nakamura, Kazusa DNA Research Institute, Department of
 Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
 Fax:81-438-32-3934)
 On Aug 9, 2000 this sequence version replaced gi:2656032.

COMMENT
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MZF18
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Graal
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Graal-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremli.zool.jastat.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is M001 and the 3' clone is M188.

FEATURES

source
 1. 65958
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="MZF18"
 /clone_11b="Mitsui P1"
 join(355..1064,1116..5684)
 /note="gene_id:MZF18.17"
 /codon_start=1
 /pseudo
 /evidence-not_experimental
 /product="non-LTR retroelement reverse transcriptase-like
 protein"
 join(8240..8326,8416..8695,8789..9161,9246..9406,
 9571..9698,9776..10024,10109..10399)
 CDS

/note="gene_id:MZF18.16"
 /codon_start=1
 /evidence-not_experimental
 /product="(+)delta-cadherine synthase (d-cadherine
 synthase)
 /protein_id="BAB08719.1"
 /db_xref="GI:9758220"
 /translation="MGSEVNRPLADFPANLWEDPILNFSKSDLGTEFTKEKSTLKEA
 VKEAFSSSKANPIENIKFIDALCRGVSVHFEDKDYEOJDKSDTCLDFQDMQEQEC
 LVTGIIIOVFPQFGPKLSADYPERKPDNGKFKGLHVTDAIGMSLDEAQMGTGE
 CDPIILAFSRSHLEISRSRSPHAIKRLNKLKPHYHGISRIETROYISYEES
 PEPQISQARVITTMALILPTALDDDBDEIPDPMKFTYVNWVFWFKLDELEECGRS
 GCGFLKLSLOKTANGYQOEAQKMLKDVIAPEDEYKENALISGGYVALIDAEFVRMT
 VAKDAEFMLSSHPRKIRVAASEIISFTSDIISYEFEKHREHYATGIDCMQOVSKE
 RAVEYMGNIIVSDAMKDLNOELMRPHVFPPLMRVNLNLSRIDVYRYRQDATYTRKLL
 KEHIVSLIETIP1
 11597..12883
 /note="gene_id:MZF18.15"
 /codon_start=1
 /evidence-not_experimental
 /product="acetyl-CoA:benzylalcohol acetyltransferase-like
 protein"
 /protein_id="BAB08720.1"
 /db_xref="GI:9758221"
 /translation="MKLELLSEKVIKPAAPNHIOFLSLFDQPLPSYVAVIPEYD
 HSNQEDIORLKSLSQTLSEFLYPLAGQIKDQVHCNDEGALFEARKEITLSDFLRN
 PSDADLQKFTVSPDHADPEWPLHVKYITFPKGRVAVASVSKICDANSLSEVC
 SMTRASKGVADTVDPPEVGNDFYPADISIEFPLVHETSKTKREYFGSLMEIKL
 NRSASSGKRVQATRIESTITALLKCMQIKAGSKSGKVEFAITOTMDLRPVSSSLP
 HKAIQNFPIPLKESSESKMEIETVSKLOKTQOENELIRNSEDKSSVEAKERI
 ASAMLSLSEIPEMTEYVAVSWCRRMSYEFQWGPVWVAPDSVDKTVLMD9KD
 SEVEARVLTPEPDMARKYEHDSLELVYATPSPSLI1"
 complement(join(13438..13580,13695..14419,14516..14753,
 14956..15283,15400..15634,15951..16167,16240..16311,
 16400..16553))
 /note="gene_id:MZF18.14"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08721.1"
 /db_xref="GI:9758222"
 /translation="MGNMRSLVKTLMVFLVGLMVLVIMISTNLFKSKMPTSKYLN
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYHDLAFDLIFPLDMSASDISRLNLIETVREDKLEPETHDRILQ
 TWKFPQPLDSPISVLPGNNEFLMGLVNLVLSVMEFLLGITRYIYVDDHTGSI
 YNEFSYGLMDMEFGSACIFISSVLMVLRKQKEDKERNVOVSVEROTSPRS
 WPHGERLEAKSVYOSIYOATSVHFGSKPLKLLLEAESDEYGVAVCCPKRMKREY
 AKICSSGLKLNHPAISFNW"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="FRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNMRSLVKTLMVFLVGLVFWIMISTNRRONTITPKLAKYLK
 TVYFGQGNVLVLTVPMMFLVAVLSCVYILIRKLSQSYECREKMKLRMGVYN
 MNPGLVITALEFSLFSLVALLAMSLNXYLSYHVLHNDNNAKIQOAFRAFGRL
 GYGVNLCMAFLFEPVTRASTIILPLVGLTSSSTIKYHMGVNSFECLVHTVFLIY
 AMINKLMEFANMPYVPLNAGTIAWVIGTAMVTSIPSTRKKFELEFTHLYGLY
 IVEYVTHVGDSSWCMITLPIELFIDRIYLFQDSTKRSRLVSAKILPSDNLELFSKT
 PGLHYNPTSLPIHVSISIKTQHPPTITSSNLEKQTSVARIRODSWQKLYTHLS
 SDLSLEVSTEGDFGNSFDVSRHNSLIVSGSGITPFTISVIRELISQONKSTKL
 DVLVCSERKHYH


```

gene
CDS
    GGBGMLCSVDNMDVIGIHWIIGLILFGIWHILIKPMAARAEVSGRAVSYSL
    GATATMGIMACCFSEFNNTAVPSEFTGPTGPEASQAQAPFLVNDQRIAGATASQCP
    TGAKYLMRSPSGEIIIGGETEMFMDRAVIEPLRGPNGIDLKRLNDIQPWERRS
    AEYTHAPILGINSVGATEINAVNVSPRSLATSHETLAFFFFFVGHIDHGRARA
    AAGFEKIGIEREPVLFMKPLD"
    complement(3568. .4629)
    /gene="psbA"
    /complement(3568. .4629)
    /codon_start=1
    /product="p1 reaction center protein of photosystem II"
    /protein_id="AAP43791.1"
    /db_xref="GI:7259490"
    /translation="MTATLERRRESANIMGRCEITSTENRLYTIGMGVIMIPCLTFA
    ISVYIIAFAAPVDIDGIREPVSGSLYGNIIISGVIPMSNAIGLHFTYMEASL
    DEMLYNGSPYLVVCHFLGACYGREMELESLFRLGMPWIAVAVSAVAATAVFLI
    YPIGOSFSDMPGIGSTENFMIVFOAHENILMHPFIMLGAGVFGSLFSAMHSL
    VTSLIRPTTENSGANGCYKRCOEETNYIVAAHGYFRLIFQYASFNSRSLHFLA
    VMPVGLMFTAMGISTAFNLNGRFENOSVYDSOCRAINTWADIINRANLMEVYHER
    NAHNPPLDASVEAPAVNG"
    complement(4754. .6550)
    /gene="ftsI"
    /complement(4754. .6550)
    /gene="ftsI"
    /codon_start=1
    /product="putative peptidoglycan synthetase"
    /protein_id="AAE43792.1"
    /db_xref="GI:7259491"
    /translation="MNRKRIFGFSRIVLWILFFSGSSLLGLRFLYOVKMGSLTNK
    AKNOOTIITLFOPRRITCDRNGIPLAIDTLAVDIFAPHLHFKSTFDIANLYSLN
    LDVEYLOWLFTKNTGICGICIAHQAPQIOLAIKANEGLIEIQQHRRYRYPQIOLAD
    IGVYNTLHEGQAGLESCQESLOSPVSAIDGRGLINDGIPRELKQSLDLOL
    TTDLDQKASTLATIDGKCKNAKRGVYLLDPYGAITATVATAPSYDNNVYDPEIE
    RKNMPVIDLEPGSTFRINNAIALEKATIKNDFDEGCIQISDPTITNNYNNK
    QPACRNSHLNTDVLSSNSVNGVHILQRLAPETIYQITQKLGNNVFLTDPLS
    SYSKLNILEFTSYNIESAVTSFGQLAMTPRIKLAQLACLSNGGNTIRPTVLDGFD
    IONEKLEFTLNNNIPQONISLKRLLKTKVTSPTQEIYDMLDEVIENRTGSSCELG
    YRIQGTGTSOKHAEOGYSSTKHLITSFAIRPINNPQVILSVIDEPSIPISFGSNT
    AGPVYRSTIESIRIKKIPSPITLHHYCK"
    complement(6788. .7477)
    /gene="clpP"
    /complement(6788. .7477)
    /gene="clpP"
    /codon_start=1
    /product="proteolytic subunit 2 of clp protease"
    /protein_id="AAE43793.1"
    /db_xref="GI:7259492"
    /translation="MPGIPKVAIRIPGEAVSYVYVYVNRIVRERTILFGEPLDDEVA
    NDLIGVAVFLNSDDTKGLFEYINSPGSMNGLGYDMLQHNVDYTTICNGLASM
    ASFIILAGTPOQRLMPHARVMLHQPMGNGCKAKYVVESEVKKRLRELIHLVYKR
    TGOSLEKIRKDMRDNMRPQAKEXGLIDHWYTNVNEIDELDKOSNDLKIKGVMLT
    NIETSNKSELK"
    7825. .9351
    /gene="psbB"
    /gene="psbB"
    7825. .9351
    /gene="psbB"
    /codon_start=1
    /product="CP47 chlorophyll a protein of photosystem II"
    /protein_id="AAE43794.1"
    /db_xref="GI:7259493"
    /translation="KGLPRTVRYTVVNLDPGRLLSYVIMHTGLVSGWAGSMARVELAV
    FDPSPDVLNPMRQGMFLPFRTRLGLISKSGWMDINDGSTITDPGLMSYGVAAHTI
    LAGIMFLASMHVYMDLEFRDPRTRGKPLADLPKIFGILFSLGCLCFEGFAGFHTG
    LFPGIIVSDPYGITGRVQPIEESWGDGDFPNPBGIAHSHIAGILGLAGLEHLS
    VRSPFRYKALRMGNVETVSSIAAVMAFVVSCTGMYSATPIELFGPTRYQMD
    LGYFNKEIKRVQASIASGSTASPMRSRDEKLAFYDITGNPAGKGLFPAAGANNND
    GIAAGLIGHAVFDKREGRRLFVRMPPEFTFPLAVYLDKQIVRADIPERRAESKYI
    EQVGVSAFTAGGELDGVTEKOPTTYKKYARRAOLGELFEEDPARLKSQGYRSPRW
    FTGHLCLFALLFEFGIHWGARTIFRFVAGIDEDLDQVDFGAFQKLSASTKQAV
    "
    9479. .9574
    /gene="psbJ"
    9479. .9574

```

```

gene
CDS
    /gene="psbH"
    /codon_start=1
    /product="p1 protein of photosystem II"
    /protein_id="AAP43795.1"
    /db_xref="GI:7259494"
    /translation="MEALVYFLLVGLGILIEFAIFRPREPRIIK"
    complement(9669. .9800)
    /gene="psbN"
    complement(9669. .9800)
    /gene="psbN"
    /codon_start=1
    /product="N protein of photosystem II"
    /protein_id="AAE43796.1"
    /db_xref="GI:7259495"
    /translation="METATIRISFSCILLIGTGYSLTYSFGNASSSLRDEFEHED"
    9891. .10112
    /gene="psbH"
    9891. .10112
    /gene="psbH"
    /codon_start=1
    /product="10 kDa phosphoprotein of photosystem II"
    /protein_id="AAE43797.1"
    /db_xref="GI:7259496"
    /translation="MADTSQGRRTVGVNFKPLNSEYGVKAPAGMGTIVLMGVMAFLA
    VFLVILQIYASVLLDGIIPANWSSLSKI"
    10445. .11092
    /gene="petB"
    10445. .11092
    /gene="petB"
    /codon_start=1
    /product="apocytochrome b6 of cytochrome b6/f complex"
    /protein_id="AAE43798.1"
    /db_xref="GI:7259497"
    /translation="MSKYVDMFNDRLEIQIADITSKYVPHNIFYCIGITFTTCE
    IMQVASGFAMTEYRPTVTEAFASVOYIMTDVNFQWLRSIKHWSASMMVLTMLHVE
    RYVILGSEFKPRELIVPVTVGIVLAVCTVSGVGYSLPMDQGVMAVKIYTVGPDAIPV
    ICAPVELLRGCGVQGSILTFRTFVLSLHTFVPLPLAVFMAFLMRKQIGISGPL"
    11202. .11684
    /gene="petD"
    11202. .11684
    /gene="petD"
    /codon_start=1
    /product="subunit IV of cytochrome b6/f complex"
    /protein_id="AAE43799.1"
    /db_xref="GI:7259498"
    /translation="MSLQKRPDLKDVYLAKLAKGCHNYGEPANNDLYTFEPCI
    LGTIGCLVGLAVLEPTMFGSEPPNPAFPTLEIDEMTFEFPYQILRIYKPKLGSVMA
    GVPAGLITVFIESINKFONPFRPAPAMTVFLIGIVAAVWLIGATPLIDTSLTGF"
    complement(11779. .12192)
    /gene="rps9"
    complement(11779. .12192)
    /gene="rps9"
    /codon_start=1
    /product="ribosomal protein s9"
    /protein_id="AAE43800.1"
    /db_xref="GI:7259499"
    /translation="MSYQKMAPIINNYIGTGRKSAVARVLVPGNGEVLINGLPNTY
    LQFNSTLSAVRSPLLETLEGDNDYIIVKRVGGIGQOAMAILGVARALCIDTSNR
    HPLKEGFLTRDSRVKRRKTYGLKAKKARQFSKR"
    complement(12245. .13195)
    /gene="rpoA"
    complement(12245. .13195)
    /gene="rpoA"
    /codon_start=1
    /product="alpha subunit of RNA polymerase"
    /protein_id="AAE43801.1"
    /db_xref="GI:7259500"
    /translation="MSGNDLFPSTIYCIESKIESPRNLGYRFLIEPLAIOGTLTVNT
    LRRLIDIGCAITTSVIGKANNEFSILPGIRESVLEIILNLKEIYFRKRSIDVQNG
    YLSTQGCVCVKANLQPLTSIEVVDGGQYIATLSGNANLDMEFVINTGKGYOMADYAI
    "
    25.38; Score 247.4; DB 6; Length 118360;

```

Query Match

NAHNPDLIAVLEPAAVNG
6178. .6300
/note="ORF40a"
/codon_start=1
/protein_id="BAA57843.1"
/db_xref="GI:2224359"
/translation="MVRCSIHETFSNFKRECVELCLQNTLYKKAVYELARSS"
complement(6212. .6379)
/note="ORF55a"
/codon_start=1
/protein_id="BAA57844.1"
/db_xref="GI:2224360"
/translation="MLAVKAYGLEERNCDLFTQIQLGVLVPLNYLSAHKQFFYIEYFAG
HIIHIIHVVN"
6803. .6943
/note="ORF46a"
/codon_start=1
/protein_id="BAA57845.1"
/db_xref="GI:2224361"
/translation="MESKQISFFIFFSFFLEFFPKKKERASPAFIPLSPFFCFYKK
SR"
7448. .7693
/gene="psac"
7448. .7693
/gene="psac"
/codon_start=1
/product="Photosystem I iron-sulfur center"
/protein_id="BAA57846.1"
/db_xref="GI:2224362"
/translation="MSHRYKYPDCIGCTGCTGACPPDYLEMPMDGCKASQASAPR
TEDCVGCKRCESACTDFLSVRYVLGGSETTSQGLAT"
7840. .7989
/note="ORF49a"
/codon_start=1
/protein_id="BAA57847.1"
/db_xref="GI:2224363"
/translation="MKELTFKFFRIASFLEKKILFFSQSKKTIISVWRKIKRTIFFE
LFCDS"
complement(7986. .8297)
/note="ORF103"
/codon_start=1
/protein_id="BAA57848.1"
/db_xref="GI:2224364"
/translation="MNVFDLILITFLYGNLFAIONCVLHNPVFIYLATVFEFNN
KILYFLQEKDVOSTRYFEFTKKKDVWRNPNFYLSINLTIVYKGFLLGFFVEAFKVG
S"
8330. .8413
/gene="trnI"
/product="tRNA-Ile(CAU)"
8330. .8413
/gene="trnI"
complement(9037. .9183)
/note="ORF48a"
/codon_start=1
/protein_id="BAA57849.1"
/db_xref="GI:2224365"
/translation="MFSKKKFSKYFILNLYDTSKKMMRFLQGFILRLIAKPFNN
KTRK"
complement(9141. .9557)
/note="ORF138"
/codon_start=1
/protein_id="BAA57850.1"
/db_xref="GI:2224366"
/translation="MENATVPNDONEDYIDLTKPMKSGYMKMKVKGFYRWVLVDFP
KINDILNTEIYPLIVSGAKTFTKSRFSMRRSLMEFAFYVCGFFLLTFYKKKKKKR
KISNTSKKRRTRKKKTRNLGNLCELRKMLASILFF"
9485. .9745
/note="ORF86b"
/codon_start=1
/protein_id="BAA57851.1"
/db_xref="GI:2224367"
/translation="MTYSVLNLCLNRGHELRHLPSEYTLQMTHHERRKRKFLFLE
FEPQDERSYKIRPKTNHTGVARLLVLRKLILKTAPKEYLHT"

Query Match	Best Local Similarity	Score	DB	Length
Matches 439; Conservative 0; Mismatches 284; Indels 15; Gaps 1.	59.5%	243.6	6	150613
Query 174	gctatcgctgcttaccctccggaagagcgctgttgaagaagacacaaccaccgcaa	233		
Db 99337	GGGTCGTAATTTGTTATTAACCTTCAGCAAGAGTGGGTTGGGAACAACAACAGCAAA	99278		
Query 234	tgtcgtctctctcgcgtcgtaagcttccagttgctgcacatgacgcgcaccttgg	293		
Db 99277	TTTAGGAAATGCAATATGCGCGCTGGTATTCGTTGCTCTTATGTAGATGACAGATATAG	99218		
Query 294	tctcgttaacctgctccctccctccttagaggttagaagatcgatcaacttaccctgcga	353		
Db 99217	ATTACGTAATTTAGATTATTTATTTAGGATTTGGAATTCGTGTACTTATACCTGCAATGA	99158		
Query 354	ggttaaaagagatgtgcctcgcatacaagctctgtagatgtaaaagcgttgcga	413		
Db 99157	TATTTGTAAGGTCAATGTCGCTCGATCGATCAACCTTTAATTCGAGATTAACGATGCAAAA	99098		
Query 414	tttgcgaattgctatgtatatctaaacctagatcgaaaactccgagtggaattggtgtaa	473		
Db 99097	TTTACCATTTATAGCAATTTCTTAAAAACGACAAAA-----TATATATGT	99053		
Query 474	agcaatggaatgctgtgagatgctgtgaaacctagaccggaaggtttacacggattcat	533		
Db 99052	TACAGAAAGAACATGCGAGAAATTTATTTATTTCTGTAAAGAAACCTGGTTTTCAGTTTG	98993		
Query 534	catcatgcatgtcctcgaagaatcgatgacgattcataaacgcacattactcggcgaa	593		
Db 98992	TTTATATGATGTCCGTCGCTGATTTGACCTAGCATTTATTAATTAATGACCATTTGACATCTGCACA	98933		
Query 594	tgaagcagttctcgtgtaacaactccgagataaaacagcglttaaggaatgctgataaggttacc	653		

gene
CDS
5489..6964
/gene="atpb"
/product="CpI beta subunit of ATP synthase"
/codon_start=1
/protein_id="AAD54782.1"
/db_xref="GI:5880689"
/translation="MSQAGVKTKQAKGIVQIVGPMYDAFQPGSMPTNYMNLIVKGRN
GAGGEVSVCEVOQLIGDGLVRAVMSATIDGIMKMEYTDGRALSYVGGTTLGRIP
NVLGEVDNMGVYNEKTLPIHREAPAFDDTLKLTSETGIKVVDLAPITREGKIG
LBSGAVPDTYIMELINIAKAGSVSVFGVGRTEGDMLYMEKESKYINKADV
KSRGALVYGGQNEPRGARMYRGITATMAEYFRVNNODVILFIDNTERVQAGEV
SALLGMPSPAVGQOPTLATMEGSGLOERTSTKDGSIQAVYVPADLITPAPATTF
AHLDATGFSRLNLAAGITIPAYDPDLSSTMLQPMILGDHDTGAORVKQTLQRYKEL
ODIALIGDELSEEDRMIVARRKIERLSQPIFYAEVGTGAPKYYLGEISIGGFK
LILSGEDSLPESAFVGNIEFATAKAASIONAAAAAA"
6977..7378
/gene="atpe"
/gene="atpe"
/gene="atpe"
/codon_start=1
/product="CpI epsilon subunit of ATP synthase"
/protein_id="AAD54783.1"
/db_xref="GI:5880690"
/translation="MSLQVRITLPEDVFLVNSADEITLPNNGMGVLTINHPILITAI
DIGMLRVSESTQSMALIGLGLAVKDNQVITILVNEAELOSDINAEAEITPLAAEA
LANSKTRDIEENILAFRRARVQVATLVK"
complement(7622..8104)
/gene="petd"
complement(7622..8104)
/gene="petd"
/codon_start=1
/product="subunit IV of cytochrome b6/f complex"
/protein_id="AAD54784.1"
/db_xref="GI:5880691"
/translation="MSVYRKPDLDTPVLRALKAGMGHNTYGEAPPNOLLVMPPVYI
LCTSLCTIGLALDPDAIGEPANFPAPLEILPMPYEPVQOLRTVPNKLTGLVLA
VAPACLIVLPFLESINKPQNPERRVATVTVLIGVVAIMWLGATLPLIDISTLTGLA"
gene
CDS
complement(8132..8779)
/gene="petb"
complement(8132..8779)
/gene="petb"
/codon_start=1
/product="apocytochrome b6 of cytochrome b6/f complex"
/protein_id="AAD54785.1"
/db_xref="GI:5880692"
/translation="MSKIYDFEERLEIOAIADITSKYPPHVNIFCYCGSITLWCF
LIVQAVGFAWTEIYRPTVTEAFASVEYIMFNINFGILRSIHMSASMMVMMLILHVF
RVYLIGFPKPRELTWGTGLAVTVSGGVYGISLPMQOVGVMAVKIVTGVPDAPIV
IGAPVLELNGSVSVSQSTLRTSLHTEVLPILTAVFMHFMIMRKQGISGPI"
complement(9045..9431)
/gene="rpl12"
complement(9045..9431)
/gene="rpl12"
/codon_start=1
/product="ribosomal protein L12"
/protein_id="AAD54786.1"
/db_xref="GI:5880693"
/translation="MSTITNQLIIBQLKSMITLLAEELVQIETPGVDASAPACAGVM
VAAAPASPAVVEQTEFTYMINDPASKRRIAVIKVVALTGLGKIKKAKMIEVPRAI
KEGIGDAQOIKOOLEEAGATVVK"
complement(9442..9846)
/gene="rps9"
complement(9442..9846)
/gene="rps9"
/codon_start=1
/product="ribosomal protein S9"
/protein_id="AAD54787.1"
/db_xref="GI:5880694"
/translation="MDHTSSGLKRTGKRCANQAVRIETGGTGLTINQRPVYLDY
NVAYLLAOGALDIWELTNRYDTLYKTIQGGGLVQCAEAIKIGLARALCSLHIKRRKAL
KPGYLTNRPLKREKKKIGLKKARAPFSKR"

gene	complement(9998..11485)	gene="trpA"	complement(9998..11485)
CDS	/gene="trpA"	/gene="trpA"	/codon_start=1
	/product="alpha subunit of RNA polymerase"	/protein_id="AAB54788.1"	
Query Match	23.0%	Score 224.6;	DB 6; Length 200799;
Best Local Similarity	58.8%;	Pred. No. 1.9e-53;	
Matches 433; Conservative	0; Mismatches 289;	Indels 15; Gaps 2;	
OY 168	aacgcgcgcatacgctcatcaacctccggaaagagcggtgttggaaagacacacac	227	
Db 167414	AGCGTGTGTCGATGATGTCATTACATCTGGCAAGAGTGGATGGGGAACCAAGCCAC	167473	
OY 228	gcgaatgtcgcgtccctcctcctcgcgcgttaagctttctcagtlgtccgcatgaacgcga	287	
Db 167474	TCCGATCTCCGCGCATCTGTATGATTCGTCGGATATCAGAGCGCTTTAAATGATGCTGA	167533	
OY 288	ccttggtccgtaacctcgatcctcctctagagttagagaaatcgatcaatlaacatctg	347	
Db 167534	CATTGGCTTACCCAACTTCTGACCTCTTACTTGGCTCGAANAATCGGCGTGTATACGCG	167593	
OY 348	cgctcaggtataaagcgcgcgattgtcgtcgtcgaacgctcgtgtcgtgataaggtgt	407	
Db 167594	CATGGAAAGTATGTAAGAGCGCAGTGGCGATTATGAAACAAGCTCTATTGCTGATTAACGTTG	167653	
OY 408	gtcgaatttcgaaattgctatgataatacctaagctagaatcgaaacttcgaatggatttg	467	
Db 167654	GAAAACTTATAGCATGCTTACCGATGTCAAAAAATTCG-----ACAACGCTACATATATGAC	167707	
OY 468	tggtgaagcatctggaaatgctgtgtgagatcggttgaaaactagaccgcgaagttcacgcga	527	
Db 167708	TCCGAAAAAATATGATGTGATGATGTCGATTCCTATCAA-----GAGCGAGGCTATCA	167758	
OY 528	ttatcatcaatcgattgtcctcgcaggaatcgatcgcgcgattcaataacgcgaattactcc	587	
Db 167759	ATACATTTCTGATTTAGCTGTCCAGCGGGGAATGACGACGAGATTGTTCACAGCCATCGCTCC	167818	
OY 588	ggcgaatgaacgaatcgtgtgataaacctccggaatataacagcgttaagagagttcgtatga	647	
Db 167819	AGCCGATTAACCCATTTCTGGTGACCATCTCTGAATTCACACGCAATTCGACGAGGAGATCG	167878	
OY 648	ggttaacgggtttgttagaatcgatcgatgaatcagaagataaagatgatttgaacaaggt	707	
Db 167879	TGTGGCAGCTCTCTCGAAGCAACATGATTTCTACATGTGCGCTTAGTAGCCAAATCGTGT	167938	
OY 708	gagaaatgatatgattaaagagagagatataagtcgaatgttagatgtgcagagaatgtt	767	
Db 167939	GGGTCCCGAGATGATTAACAATAATGACATGATGTCTGTGATGTGTCACAGATGAT	167998	
OY 768	gggaatgtatcattgtgtgtgataaatttccttgaagaattcgaagttatcgaagcaaatgc	827	
Db 167999	TGGCGTGCAATGTTAGGCGCCATTCGCGAAGATTAAGATGTGATTTATTTTCGACCAACG	168058	
OY 828	aagggttcgcgtgttctgaataagcctcctcgaatcgttcggaatttgagcgttgaagcagc	887	
Db 168059	AGCGGAACCTTTATGTGTGTCAAAAGACATTTACGTGCGAGCGAGGTGCTCTTGAAAGAC	168118	
OY 888	ggcttgaagactcgttg 904		
Db 168119	AGCTCGACGCTTGCTAG 168135		
RESULT 7			
AF137379/c	AF137379	200799 bp	DNA circular pln 14-SEP-1999
LOCUS	Nephrolepis olivacea	chloroplast DNA, complete genome.	
DEFINITION	AF137379		
ACCESSION	AF137379.1	GI:5880684	
VERSION			
KEYWORDS			
SOURCE	Nephrolepis olivacea.		

ORGANISM

Chloroplast Nephroselmis olivacea
Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
Chlorodendraceae; Chlorodendraceae; Nephroselmis.

REFERENCE

1 (bases 1 to 200799)
Turmel, M., Ollis, C. and Lemieux, C.

AUTHORS

The complete chloroplast DNA sequence of the green alga
Nephroselmis olivacea: Insights into the architecture of ancestral
chloroplast genomes

Proc. Natl. Acad. Sci. U.S.A. 96 (18), 10248-10253 (1999)

JOURNAL

2 (bases 1 to 200799)
Turmel, M., Ollis, C. and Lemieux, C.

AUTHORS

Submitted (24-MAR-1999) Biochimie, Université Laval, Pavillon
Charles-Eugene Marchand, Quebec G1K 7P4, Canada

JOURNAL

FEATURES

source

1..200799

/organism="Nephroselmis olivacea"

/organelle="plastid:chloroplast"

/db_xref="taxon:31312"

/cell_line="NIES 484"

1..92126

/note="Region: large single-copy region (LSC)"

55..137

/gene="trnL(gag)"

/note="codons recognized: CUY"

/product="trnL-leu"

/anticodon="(pos:88..90,aa:Leu)"

55..137

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

332..634

/gene="trnL(gag)"

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

/product="trnL-Met"

/anticodon="(pos:1652..1654,aa:Met)"

1618..1691

/gene="trnM(cau)"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

/gene="trnI"

1733..3862

Cryptomonas phi does not involve a signal peptide

JOURNAL MEDLINE 6 (bases 110917 to 113854)
REFERENCE DOUGLAS,S.E.
AUTHORS Unusual organization of a ribosomal protein operon in the plastid genome of Cryptomonas phi: evolutionary considerations
JOURNAL MEDLINE 7 (bases 40675 to 42376)
REFERENCE DOUGLAS,S.E. and Turner,S.
AUTHORS Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor
JOURNAL MEDLINE 8 (bases 96129 to 98906)
REFERENCE WANG,S.L. and Liu,X.Q.
AUTHORS The plastid genome of Cryptomonas phi encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein
JOURNAL MEDLINE 9 (bases 106789 to 108216)
REFERENCE DOUGLAS,S.E.
AUTHORS A sect homologue is found in the plastid genome of Cryptomonas phi
JOURNAL MEDLINE 10 (bases 42198 to 44153)
REFERENCE DOUGLAS,S.E. and Reith,M.F.
AUTHORS A bc1l homolog, encoding a subunit of Mg chelatase, is located on the plastid genomes of red and cryptomonad algae
JOURNAL MEDLINE 11 (bases 8327 to 84479)
REFERENCE DOUGLAS,S.E. and Murphy,C.A.
AUTHORS Structural, transcriptional and phylogenetic analyses of the atpB gene cluster from the plastid of Cryptomonas F (Cryophyceae)
JOURNAL MEDLINE 12 (bases 98901 to 114602)
REFERENCE WANG,S.L., Liu,X.Q. and Douglas,S.E.
AUTHORS The large ribosomal protein gene cluster of a cryptomonad plastid: gene organization, sequence and evolutionary implications
JOURNAL MEDLINE 13 (bases 61067 to 68605)
REFERENCE LEITSCH,C.E.W., Kowallik,K.V. and Douglas,S.E.
AUTHORS The atpA gene cluster of a cryptomonad, Guillardia theta: A piece in the puzzle of chloroplast genome development
JOURNAL MEDLINE 14 (bases 1 to 121524)
REFERENCE DOUGLAS,S.E. and Penny,S.L.
AUTHORS The plastid genome of the cryptophyte alga, Guillardia theta: complete sequence and conserved syntenic groups confirm its common ancestry with red algae
JOURNAL MEDLINE 15 (bases 1 to 121524)
REFERENCE DOUGLAS,S.E.
AUTHORS Direct Submission
JOURNAL MEDLINE Submitted (08-JAN-1998) Institute for Marine Biosciences, National Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada
COMMENT On Sep 15, 1998 this sequence version replaced g1:11396 g1:11297 g1:18103 g1:18281 g1:11363 g1:11407 g1:12359 g1:136730 g1:11300 g1:398949 g1:2661180.
FEATURES
SOURCE location/Qualifiers
1. .121524
/organism="Guillardia theta"
/organism="plastid:chloroplast"
/db_xref="taxon:55529"
/note="plastid"
complement(5..86)
/gene="trnI(gua)"
/note="codon recognized: CAU"
/product="tRNA-Tyr"

gene
complement(5..86)
/gene="trnI(gua)"
complement(181..573)
/gene="rpl19"
complement(181..573)
/gene="rpl19"
/codon_start=1
/product="ribosomal protein L19"
/protein_id="AAC35594.1"
/db_xref="GI:3602933"
/translation="METILHLYFEVNNKSNKLNLIQVESNHNSTLPOLQGVDFVRLGVSITGSKRIQATGVIISKNSGIMLITITIRRVLOGIGIERVYLLHSPKLSIEILRSKIRRAKLYLRTSGKARLQORFN"

gene
729..3188
/gene="clpC"
729..3188
/gene="clpC"
/codon_start=1
/product="Clp protease ATP binding subunit"
/protein_id="AAC35594.1"
/db_xref="GI:3602934"
/translation="MEPEETKAKIVIMLAOEARRLGHNFVGTETQILIGIGETGIAAKVLSMGVNLKDAKREVERKIGRGGFVAVEIPFPRAKRVLELSSEERQGHNYIGTEHLLGLIREGEVAARVLENLADLTKRTOVRLGDTAEVSTNCGTKPTLIEFGSNLTOKAAEGKLDVIGKIEIERYIOIGRTKNNPLIGEPGQVIAEGLAQRINRDVDPILDEKRVVTLIDGLVAGTKYRGSEPRIRKITTDEIRVANNVILIDEVHTLIGAGAEADANLIIPALAREMOTCATTEETRKHIEKDSALRRQPVWGEPSVETETIILYGLDPRYKELHKKVITSDALSAKFAOVYADRLPKAIDLDEASRPKINSQPLPARBELDEKREILKODEKAVRSODEFTAGQLRDEMELHGRITGODEAVAVASAKIRRAVGLKNNRPIASIFSGPVGKTELTAKLASFPFS EEMVRLDSEMERHNVSKLISPPGVVNGGQLTESVRRPYVAVLDFDEKHPVFNVLILILEDRGLTDSKGRVDFKNTLLITSNVGSKVIEKGGGLDEQHEQOYGRITAKLVNELKQYRPFELNRIDEITVROLTKDVEGTAETMLKEVFRIS EKGIOETVTRKFTHLINEGVNPYTGAPLRRAVRLLEDTLSEETLAKIKEGTAVYDVDDDKRVKVLDEKLELVAN"

trnA
3316..3388
/gene="trnI(gaa)"
/note="codon recognized: CUU"
/product="tRNA-Phe"
/anticodon="(pos:3349..3351,aa:Phe)
3316..3388
/gene="trnI(gaa)"
complement(3402..5201)
/gene="dnaB"
complement(3402..5201)
/gene="dnaB"
/codon_start=1
/transl_except="(pos:complement(5199..5201),aa:Met)
/product="replication helicase subunit"
/protein_id="AAC35596.1"
/db_xref="GI:3602935"
/translation="MKVTKFYMLKAKENHLLIQLVEKDFLQIYTKIDTQVFEFQSHQILYRSINILITKRVLINPNSLLSHSELSERVIVYLTFLFDRNPKNRKLINOSLQVLDFNIRELONSCAKIVSTFENFNLSEVTLQKSLITSNNTNKKVAVSLRNVLDTLEIDKTKNRSTHYLGFEDLDHVLVGLQKSLITIAQSPMKVAPMLSLVRNVLADIOSFPVIVSLMSKOLYRILISNTNATSLTREGNISMETLNRMATYLSNINYLDDENNLDVLDIOSKLASIOIYDIDICVLIQOLQYKRSORSRTELSITRYLTIKADENPLPLVLSOLSRNVEFRNKRPILVSLDESGLSATQNPFLKSDNVISVETVSLKSVNHYILSTYELLPQNRITPYVSLIKREIQGTQYKVVYELDFYISICMNRINVDYETPKNPFVVALLVHNSIEDADVLEFLVRSYVQHMLKOSNIDMC ETVAKHRRGTGMVNLJFNPVTSFPMLIKES"

misc_feature
3594..4073
/note="signal element intein: encodes putative intein spliced out of the replication helicase subunit"
5361..5433
/gene="trnI(cca)"
/note="codon recognized: GGU"
/product="tRNA-Trp"
/anticodon="(pos:5394..5396,aa:Trp)
5361..5433

trnA
complement(5..86)
/gene="trnI(gua)"
complement(181..573)
/gene="rpl19"
complement(181..573)
/gene="rpl19"
/codon_start=1
/product="ribosomal protein L19"
/protein_id="AAC35594.1"
/db_xref="GI:3602933"
/translation="METILHLYFEVNNKSNKLNLIQVESNHNSTLPOLQGVDFVRLGVSITGSKRIQATGVIISKNSGIMLITITIRRVLOGIGIERVYLLHSPKLSIEILRSKIRRAKLYLRTSGKARLQORFN"

gene
complement(5..86)
/gene="trnI(gua)"
complement(181..573)
/gene="rpl19"
complement(181..573)
/gene="rpl19"
/codon_start=1
/product="ribosomal protein L19"
/protein_id="AAC35594.1"
/db_xref="GI:3602933"
/translation="METILHLYFEVNNKSNKLNLIQVESNHNSTLPOLQGVDFVRLGVSITGSKRIQATGVIISKNSGIMLITITIRRVLOGIGIERVYLLHSPKLSIEILRSKIRRAKLYLRTSGKARLQORFN"

OY	352	gaggttataaaggagatctgttcgcgatacaagctctggtaacgtttgcgc	411
Db	9811	GAATATTATTGAAGGTCGTTGTCGACTAGATCAAGCGCTGGTACGTGAAGCGCTTGGA	9752
OY	412	aattcgaatlgctalgtalatactaaactagatcgaacatlcgcgalg9gatttgctg	471
Db	9751	AATTATGCAATTATWGCGCACTTTCAA-----AAACACCCAAAATATATGTAC	9702
OY	472	aaagcatggaatgctctgtgatacgtttgaaactagaccggaaggtttacccgatttc	531
Db	9701	GCACACACATATGCGCAATTAGTTTTCATATMAAGAACTAGTAT-----CAATTC	9647
OY	532	atcacatcgcattgtccctcaggaatcgcgataccggaattcaaccgcattacccg	591
Db	9646	ATATTATATGATTGCCCCGCGAGGTATGATGTTTGTTTATTAATGCAATAGCCACG	9587
OY	592	aatgaagcagcttctgttaacaactccggaatalacagcgcttlaaggaatgtctga	651
Db	9586	CAAGAAGCAATATTGTTCACACACCTGAAATTAACGCTTTGCTGATGCGAATAGAG	9527
OY	652	accggatctgtcgaatgcgatcggaatcagagatataaagatgatttgaaacga	711
Db	9526	GCTGATATTATGGAAGCGATACAAATGTGTGAACACAAAATATTATTAATTCG	9467
OY	712	actgatatgaattaaaggagagatatgaatgcgttgatgaatctgcgagagatlg	771
Db	9466	ATGGATATGATTCCTCAAAATATGTCACAAATGTTATTCATATATGATCTTC	9407
OY	772	ttgcatatgcttgctgtaactctcgtagaaatlcgaagtlctcgaagcagaacat	831
Db	9406	ATTCCATATTATGCGGCATATACCTGGAAGATACAAATGTATTTATTCAC	9347
OY	832	ttccgccttgctcgaataagccctcgaacgcttcgaggatcgttcgatttgacg	891
Db	9346	GAACCTTTAGCTTGATGATATAAAAAATTAACGCTTTCAGGTATTCATTTG	9287
OY	892	tggagacatcgctgacga	911
Db	9286	CGTACATTAATAGCAAAAGA	9267

LOCUS	SYNGLG	135638 bp	DNA	BC1	13-FEB-1999			
DEFINITION	Synechocystis sp. PCC6803 complete genome, 24/27, 3002966-3138603.							
ACCESSION	D64005	AB001339						
VERSION	D64005.1	GI:1001779						
KEYWORDS	UmcC; ABC transporter; DNA polymerase I; P-methylase; Pld; UmcC protein; beta-lactamase; bromoperoxidase; cadmium-transporting ATPase; cation efflux system protein; cation-transporting ATPase Pcd; cytochrome b6-f complex subunit Pctm; eukaryotic protein kinase; glutamate dehydrogenase (NADP+); heterocyst to vegetative cell connection protein; magnesium and cobalt transport protein; mercupic resistance operon regulatory protein; methionine aminopeptidase; molybdate-binding periplasmic protein; nickel resistance; nitrate/nitrite response regulator protein; nitrilase; phospho-N-acetylglutamate-5-pentapeptide-transferase; phosphoenolpyruvate synthase; phosphofructokinase; polyphosphate kinase; pre-B cell enhancing factor; protein kinase; pyridoxal phosphate biosynthetic protein PdxA; pyruvate oxidoreductase; regulatory components of sensory transduction system; salt-stress induced hydrophobic peptide; sensory transduction histidine kinase septum site-determining protein Minc; septum site-determining protein Mmd; septum site-determining protein Mnm; transposase repressor SmuB; transfer RNA guanine transglycosylase; transposase Yc135.							
SOURCE	Synechocystis sp. (strain:PCC6803) DNA.							
ORGANISM	Synechocystis sp.							
REFERENCE	Bacteria: Cyanobacteria; Chroococcales; Synechocystis.							
AUTHORS	1 (sites) Kaneko,T., Tanaka,A., Sato,S., Kotani,H., Sazuka,T., Miyajima,N., Sugita,M. and Tabata,S.							

TITLE	JOURNAL MEDLINE REFERENCE	TITLE	JOURNAL MEDLINE COMMENT	FEATURES SOURCE
Sequence analysis of the genome of the unicellular cyanobacterium <i>Synechocystis</i> sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome	JOURNAL MEDLINE REFERENCE	Sequence analysis of the genome of the unicellular cyanobacterium <i>Synechocystis</i> sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions	JOURNAL MEDLINE COMMENT	Potential protein coding regions were assigned on the basis of similarity search of the ORFs and Genemark analysis.
2 (bases 1 to 135638)	96127529	2 (bases 1 to 135638)	96127529	Location/Qualifiers
Kaneke, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosewa, M., Sugiyura, M., Sasanoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Mureki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpou, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S.		Sequence analysis of the genome of the unicellular cyanobacterium <i>Synechocystis</i> sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions		1. .135638
DNA Res. 2 (4), 153-166 (1995)		DNA Res. 3 (3), 109-136 (1996)		/organism="Synechocystis sp."
				/strain="PCC6803"
				/db_xref="taxon:1143"
				1. .1488
				/note="ORF_ID:slr0121"
				/codon_start=1
				/transl_table=11
				/product="hypothetical protein"
				/protein_id="BAA10659.1"
				/db_xref="GI:1001780"
				/translation="MTSMRPRRSRTQRRNLRYISGTSIPPGVTIAEYEDGAGAS
				TEINQESNOKPGETKVNPLKLVVSATGVKSVAKGNGLRPKRSKPLSMWOKSVLVS
				VRIGVIGIGIAGIAGTAITTFSPKIFRCQNGNPLATAEHPHOTDPOAIVQKVLV
				ASLAGVATETDKRGTPATPEATPEATSEQETVATASASVDQPPQEDALOKLTL
				TLGSKGPAINSYSEFIDVNGCPAANKGEGTLPASTIKIPVAIAEFAEDGOKIOLD
				HEELPLQDIVIGVEAGTMOYDVKRSSYPLAEVYTKMIVISDNTATMLTKRGKREE
				LNRFQAMQMPQTKINNYLPDLDECTNTTSQDIALILLYKIQGGELSLSKDRLLGTM
				QQTIRLRLPQGLEPDAIOLNHRKTDIGTIVGDGIVDMPGKRTVGAVALTRPHNDVA
				GRLLIDISRTAYQHYKGIIPPASSPAPQAKESPATAVEPSEN"
				complement (1489. .3675)
				/gene="ppk"
				complement (1489. .3675)
				/gene="ppk"
				/note="ORF_ID:slr0290"
				/codon_start=1
				/transl_table=11
				/product="polyphosphate kinase"
				/protein_id="BAA10660.1"
				/db_xref="GI:1001781"
				/translation="MTQAMSPAPVDPKDDPAAYFNRELSTLAFNORVLIHGIDRTPIL
				LEBKPLAIFCSNIDPEFMVRAVGLAKQVENVATKLTADGTRTSOOLKETSKSLRPLV
				YQONQVFEYVLEKELADEGIFLNDYVDLSOEKRYLHOFYDDHLEFVLTPLVADYSH
				EPYISNLSLNGVVRDPDDELEFAKVKVPTLPREVALPEVQDOPDKPKPMLMTGVS
				PLEQVIAHNHLASLFPGMIIOECHLFTFRNADIAVEDEADDLIAIEELRKRRVGK
				SAVRLLENASTPKNIRLMTDLGLEIDYDIDGLGKDLDFEFLSLPHLKDDEPM
				ASVIRPELKHVEFEVDGEDGRVOGEEDIDLETLIRGGDILVHHYQSFATSVQOFTQ
				AAADPVHITIKMLTIRSGDSPYLAALAAENGKQAVILVELKARDEENIMAKK
				LEBOYCHVYIGVLGKTHRTVLYVROEGDPDIRRYHIGNTINPKAKLTDTDLGLIT
				CRBELGNDLNLNENFTLGYRSROKDYOKLVAAPVMRRERVAAMIERADHCLNGTGT
				VAMKNSLVDTQIIRALYVMSQAQGVQDILVIGICLCPREVENASENRTVSYIGRLLE
				HSISITFYHNGEEREIYIGSADMMSRNTRREVAAPVEQDPLKQELOSIGILLADR
				QAMLEIDGTYVORRPASPEOSOSQAFVQAIAETTEDPELRSV"
				complement (4315. .4608)
				/gene="mine"
				complement (4315. .4608)
				/gene="mine"
				/note="ORF_ID:ssl0546"
				/codon_start=1
				/transl_table=11
				/product="septum site-determining protein mine"
				/protein_id="BAA10661.1"
				/db_xref="GI:1001782"
				/translation="MIELIERLFRSNGKNSGEDARRRLVIANDRSGLSPENMEEM

gene RREIVEVRSRYVEIDPGEMFESLSLDQRTALIANLPVRHVRRTKAKSEAQES
complement(4666,.5466)
/gene="mind"
CDS complement(4666,.5466)

[illegible]

```

/ab_xref="GI:1001784"
/translaton="METNRRHOTAAASAGLSVDOTMAKPLVDPSEETSLPKPLVYR
HTLRSGGELRHGGDVTTCIDVNGGSIIVADGDILLGCLKGMNAAGAKGDAQVAIL
RLAACQITRISDRPLVAGADAVDRREPELAYITSEGIRLTPVRFORSALL"
complement (6028...5285)
/notes="ORF

```

```

6497. 7129
/gene="frah"
6497. 7129
/gene="frah"
/note="ORF_ID:slr0298"
/codon_start=1

```

UCASLVHVOITTELEIPAGLIVILGKPNKVPDPDIVSGCANDVVSRYHADRILEG
NEYFLDEYSCNGSYVNHMTALLPGNRHLIRPGRLAIGKGLVFLIQLH"
7191. .7898
/note="ORF_ID:slr0299"
/codon_start=1
/transl_table=11

```
7874...8948...17855
/note="ORF_ID:s1r0300"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
```

```
/translation="MYSTIKLRPIISLTQQFVALCQONPDLKFERSSNGDLIMP
```


Query Match 13.7%; Score 133.6; DB 6; Length 673;
 Best Local Similarity 59.9%; Pred. No. 1.5e-27;
 Matches 242; Conservative 0; Mismatches 159; Indels 3; Gaps 1;

OY 151 ccggaactcgcgggaaagacgcgcatatcgttatacaccctcgggaaagcggtgtt 210
 DB 453 cctgacctcccgccgacaccccgccctgctgctgctcactccggcaagggcgctg 394
 OY 211 ggaagaacgacacacacgcaaatgtctgtctctcctcgtcgttaagttcagtt 270
 DB 393 ggttaagacacacacacgcaaatgtctgtctcctcgtcgttaagttcagtt 334
 OY 271 gtcgcaatgaacgcaacgttgcctcgttaacgtcgtcctcctcctaggttagaagt 330
 DB 333 gtcgcaatgaacgcaacgttgcctcgttaacgtcgtcctcctcctaggttagaagt 274
 OY 331 cgaagtaatacaactcgtcgtcgttaatacaaggaatgtcgtcgtcgtcgtcgtcgt 390
 DB 273 cgcgctccacacacacacgcaaatgtctgtctcgtcgtcgtcgtcgtcgtcgtcgt 214
 OY 391 gtaagtgataagcgttgcgttaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 450
 DB 213 gtcgcaatgaacgcaacgttgcctcgttaacgtcgtcctcctcctcctcctcctcct 154
 OY 451 cttcgaatgaacgcaacgttgcctcgttaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 510
 DB 153 cttcgaatgaacgcaacgttgcctcgttaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 97
 OY 511 ccggaagtgataacgcaacgttgcctcgttaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 554
 DB 96 gccgcaacacacacacgcaaatgtctgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 53

RESULT 13
 AE000712/c 14103 bp DNA BCT 25-MAR-1998
 LOCUS Aquifex aeolicus section 44 of 109 of the complete genome.
 DEFINITION AE000712 AE000657
 ACCESSION AE000712.1 GI:2983411
 VERSION
 KEYWORDS
 SOURCE Aquifex aeolicus.
 ORGANISM Aquifex aeolicus.
 Bacteria; Aquificales; Aquificaceae; Aquifex.

REFERENCE 1 (bases 1 to 14103)
 AUTHORS Decker, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L.,
 Graham, D. E., Overbeek, R., Sneed, M. A., Keller, M., Aujay, M.,
 Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V.
 TITLE The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus
 JOURNAL Nature 392 (6674), 353-358 (1998)
 MEDLINE 98196666
 AUTHORS

2 (bases 1 to 14103)
 Decker, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L.,
 Graham, D. E., Overbeek, R., Sneed, M. A., Keller, M., Aujay, M.,
 Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J., and Swanson, R. V.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-1997) Diversa Corporation, Genomics, San Diego,
 CA 92121
 COMMENT Putative indicates no similarity to known proteins
 Hypothetical indicates similarity to a protein of unknown function.
 FEATURES
 Location/Qualifiers
 1..14103
 /organism="Aquifex aeolicus"
 /strain="vps"
 /db_xref="taxon:63363"
 /complement(150..938)
 /gene="mind2"
 /note="ag.877"
 /complement(150..938)
 /gene="mind2"
 /codon_start=1

gene
 CDS

1..14103
 /organism="Aquifex aeolicus"
 /strain="vps"
 /db_xref="taxon:63363"
 /complement(150..938)
 /gene="mind2"
 /note="ag.877"
 /complement(150..938)
 /gene="mind2"
 /codon_start=1

/transl_table=11
 /product="septum site-determining protein Mind"
 /protein_id="AAC06996.1"
 /db_xref="GI:2983413"
 /translation="MAEYVITSGGCGVKTLLTNIGALACKKVLIDADIGLR
 NDMILGENTVYDIDVLSGRVPEKALVNDKRGSLIMLPAQORAKDVIDIEK
 NKTVEIKNSGNYDILVDSFAGIEKQIVASPADKALIVNPEVSSIRDDRVYGL
 LESMDKRNKYIVNRIKWMKRGALMVEDVIDILKKEIGIIEEPKLVPTNNGE
 PIVDERFPASQALIDTPARRKMGESIPIKRGEKGLLSRLIGS"
 complement(917..1522)
 /gene="mind"
 /note="ag.878"
 /complement(917..1522)
 /gene="mind"
 /codon_start=1
 /transl_table=11
 /product="septum site-determining protein Minc"
 /protein_id="AAC06995.1"
 /db_xref="GI:2983412"
 /translation="MIEIKGKTLPIYQIKREKNDLKLKELKSHINIKGSLI
 LENEVLKPERKVEELIKERSRGTEKRGKREKRESRLIIEPTLACQRIEHR
 GDILILGDNKDAEVLAGNIIYMKLGVAKGLICHSAIYVALKMPQLQIGKK
 KAINSEADRNPGEVAKIEGSDIVLEPIEGAKEMKLIL"
 1546..2124
 /gene="ag.880"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAC07003.1"
 /db_xref="GI:2983420"
 /translation="MDVFLDTSVFTNPENIYRTPEEDORGAMEFTHIALNRAEEM
 PTSYTEMKRLMDYGELEMEFEEMVYKIRSPFOLITADPLEEELERLYNNGR
 IAEHTRASCEPDYKGLIARLREKRRALNGIILDSKEDVDYLLAYELDGVLSAD
 EGLTMDKTIKILIDPKRNKILNLESLVHRP"
 complement(2121..4007)
 /gene="ag.881"
 /complement(2121..4007)
 /gene="ag.881"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAC07004.1"
 /db_xref="GI:2983421"
 /translation="MLEKYLKDKKGGPDIKNYDETLOKLAQEVROYIINWTSKNG
 GHVPSLQVVELTIALLRVNPPEYITWYDIGHQYPAKILDRKEQPTLLKQYGIS
 GLNRRESIDAFAGHSTISIALGRIGKDLKGEEDYVIAVIGGALTAGMAVE
 ALNNAGHILRPDRFIVLINDNEMISIPNGAISTYLNRIISGFVQETROKIKNFQHF
 GEPFLRIKMLTEEFELKGLISGVIEELFVYIGPIDCHDKALDELTNNVDIGPV
 LHVYTKRGKGYKPAENPVMKGVAPYKVSGETIKRSSPTMTSVGKALVELAER
 DEKIVATTPAMRSGSIVFEPAKRPDRPDYGIAGQACHPAAGLAAGLRVAAVYS
 TFLORATDOVTHVVALOMLVYTPAIDRAGLVGDSPTHHGVFDLSYLKCVPMNVYCAP
 KDELDELDLTLTISYSEKPFALRYPRGAAYVPEEGFKIEIGWEELLEGDCVILA
 VGVPYVALRAERLYKXGIVGVNARFVPMDEKMLRDLANRDFTITVEDNTVVG
 GGVSGVLEEFARIEIMKRVIMLGVDPRIEHRKODIILNLVIGDAEGIEKAVRALKG
 GRLT"

complement(4445..7885)
 /gene="top62"
 /complement(4445..7885)
 /gene="ag.886"
 /note="ag.886"
 /complement(4445..7885)
 /gene="top62"
 /codon_start=1
 /transl_table=11
 /product="reverse gyrase"
 /protein_id="AAC07000.1"
 /db_xref="GI:2983417"
 /translation="MIKAFDTLCPNGGEISABRLKGLPCQKLPPEVNBREYCOK
 LENSFKEFCLELSELDWDEFEKELIGTSWSIOKSAARVPIGRSAMLAPQVCK
 TTRGLMASYLAQKQKSYIILPOLLYEVSERKRTGVEYEDRLITWQKLSKKKE
 LKERIQKDPDILITISFLKNTYELLKPKDSITFVDVDSFLKTAKNVDKVLILGF
 SEEDINAFELIRLEKRPNKSEDEWEEIKRSEDEIETLAKKKKGLAVSSAAGNPSRN
 RIKLRELIGFEVGPVSLNRIYETETONLKETLLKRVKEFGKGLVFEVSSDGR

BSM1NCD 1900 bp DNA BCT 02-MAR-1993
 LOCUS BSM1NCD
 DEFINITION B. subtilis mnc, mind, and spoIVFA (partial) genes.
 ACCESSION 215113
 VERSION 215113.1 GI:49307
 KEYWORDS mnc gene; mind gene; spoIVFA gene; sporulation operon.
 SOURCE Bacillus subtilis.
 ORGANISM Bacillus subtilis.
 Bacteria; Firmicutes; Bacillus/Clostridium group;
 Bacillus/Staphylococcus group; Bacillus.
 1 (bases 1704 to 1900)
 Cutting, S., Roels, S. and Losick, R.
 Sporulation operon spoIV and the characterization of mutations
 that uncouple mother cell from forespore gene expression in
 Bacillus subtilis
 J. Mol. Biol. 221 (4), 1237-1256 (1991)
 JOURNAL 92046062
 MEDLINE 2 (bases 1 to 1900)
 REFERENCE Lee, S. and Price, C.W.
 The mncD locus of Bacillus subtilis lacks the mnc determinant
 that provides topological specificity to cell division
 J. Microbiol. 7 (4), 601-610 (1993)
 JOURNAL 93211302
 MEDLINE 3 (bases 1 to 1900)
 REFERENCE Price, C.W.
 Direct Submission
 Submitted (15-SEP-1992) Chester W. Price, Department of Food
 Science, University of California, Davis, Department of Food
 Science, University of California, Davis, California, 95616, USA
 TITLE Location/Qualifiers
 1. 1900
 /organism="Bacillus subtilis"
 /strain="168"
 /db_xref="taxon:1423"
 /clone="psv4, psv13"
 173. 180
 /note="putative sigma-H recognition sequence"
 194. 198
 /note="putative sigma-H recognition sequence"
 212. 218
 /gene="mnc"
 join(212..218, 224..904)
 /gene="mnc"
 224..904
 /gene="mnc"
 /function="Cell division inhibitor"
 /codon_start=1
 /transl_table=11
 /product="Minc"
 /protein_id="CAA78817.1"
 /db_xref="GI:580892"
 /translation="MKTAKQAVYVITGTNKLTLHLDDACSPDELLDGLNMLSEIY
 TDGKGKISVHVAKLGRPLKREDEQTEILASKKDLFVSISSSEVTKKEAOIREE
 AVIISVKIPIVSGQVLOVKGDLILIGVNGGIVRGGNIFVIGSLKGIHAAGNGNN
 QAVIASSEMLPQLRIHVLRSPDHIQKGNEMECAYLDTGNNVIERLQHLALRPD
 LRLREGGM"
 893..897
 /gene="mind"
 893..897
 /gene="mind"
 906..1712
 /gene="mind"
 906..1712
 /gene="mind"
 /function="ATPase activator of Minc"
 /codon_start=1
 /transl_table=11
 /product="Mind"
 /protein_id="CAA78818.1"
 /db_xref="GI:580893"
 /translation="MGEBAIVITSGKGVGKTTTSANLGTALALIGKRVCLVDTIGLR
 CDS
 RBS
 gene
 gene
 CDS

LDVVMGLNRIIYLDVVEGRCKMHAQVKKRFPDDLTYLMPAOTSDKTAPEQ
 IKNNVDELKOEEDVVIDCPAGIEOGKKNVAGADKATVTTPEISAVRDAIDTIGL
 FORENVEPEPLVVRIRIHNLMKNGDMDIDEIVHLSIDILGIYADDEVIKASHGE
 PIAMDPKNRASIAIRNIRARLIGSVPLQVLEQNKMAIKSFVGRS"
 terminator
 1737..1769
 1862..1900
 /gene="spoIVFA"
 /gene="spoIVFA"
 1862..21900
 /gene="spoIVFA"
 /function="coupling gene expression"
 /citation=11
 /codon_start=1
 /transl_table=11
 /product="SpoIVFA"
 /protein_id="CAA78819.1"
 /db_xref="GI:49310"
 /db_xref="SWISS-PROT:P26936"
 /translation="MSHRADEIRKLE"
 BASE COUNT 600 a 341 c 445 g 514 t
 ORIGIN
 Query Match 13.2% Score 128.8; DB 2; Length 1900;
 Best Local Similarity 53.0%; Freq. No. 4.3e-26;
 Matches 355; Conservative 0; Mismatches 297; Indels 18; Gaps 3;
 QY 182 tcgtatcacctccggaagagcggtgtgtgaagagcagcaaccacgcgaatgtcgtc 241
 DB 919 TCGTAAATTAACCTCGGGAAGCCGAGTAGTACGACACACATCTCGCAACCTCGGTA 978
 QY 242 tctctcgcctcgttaaggttctcagttgtcgcacatgaacgcacattgtctcgt 301
 DB 979 CCGCCTTAGCCATTAGGGAAGCGGTATGCTTAGTATGATCTGATATAGAGACTCGCA 1038
 QY 302 acctcgaatcctcctcctaggttagagagatcgaatcgaatcgtcgtcgaatgaat 361
 DB 1039 ACCTGATGTTGTAATAGGCTCTTGAATAATGATTTATTACGATCTGCTGTAAGACTGTAG 1098
 QY 362 acgagatgtcgtcgtcgtcgaatcgaatcgtcgtcgtcgtcgtcgtcgtcgtcgtc 421
 DB 1099 AGGCGAATGCAAAATGATGACGCGCTCGTAAAGACAAACGTTTCGATGATCTGCT 1158
 QY 422 tgcattgatatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 481
 DB 1159 ATTTA---ATGCCCGCTGCTCAAAACGATTAAGACACTGTTGCTCTGAAACAATT 1215
 QY 482 aatgctgtgtgattgattgaaactagaacggaagttcaacggaattcatcatcag 541
 DB 1216 AAAACATGCTCCAGAGCTCA-----AACAGGAATTTGACTATGTCATCATAG 1263
 QY 542 attgtcctcgaagatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 601
 DB 1264 ACTGTCCTCCGCAATCGACGAAGGGGTACAAAATCCCGTTCCGAGCTGATTAAGCGA 1323
 QY 602 tctcgttaacactcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 661
 DB 1324 TTGTGCTACTACGCTGAAATCTCAGCTGTTCTGATGCTGACCGTATTTAGGACTGC 1383
 QY 662 tagaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 718
 DB 1384 TGGAGCAAGAGGAATAATTTGAACCGCGCGCTGCTTTTAACAGAAATCAGAAATCACC 1443
 QY 719 tgattaaagagagatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 778
 DB 1444 TGATGAAAAACGCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
 QY 779 tgcctggtgattcctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 838
 DB 1504 TCGTCGGAATCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
 QY 839 ttgtctga 848
 DB 1564 TTGCGATGGA 1573

Db	3745	ATTTA---ATGCGCCGCTGCTCAAAACAGAGCATTAAGACAGCTGTTGCTTCGACAAATA	3801
Oy	482	aatgcctctgtgatatcgcttgtaaaactagaccggaagttaccggatttcaatc	541
Db	3802	AAAACATGGTCCAAAGGCTCA-----AACAGAAATTTGACTATGATCATCTAG	3849
Oy	542	attgtccctcagaatacgaatgcggatctcaacccgccttactccggcgataaagcag	601
Db	3850	ACTGTCCTCCCGGAATCGAGCAAGGTTACAAAATGCGCTTCCGGAGCTGATTAACCA	3909
Oy	602	ttctggtaacaactccggaataatacgcgltaaaggatctgataaggtacggtttgt	661
Db	3910	TGTGGTCACTACGGCTGAATCATCTCAGCTGTTGCTGATGCTGACCGGATTAATFAGAGATGC	3969
Oy	662	tagaatcgcgatggaatcagaga---tataaagatgtattgaaacagatgaaactgata	718
Db	3970	TGGAGCMAAGAGAAATATTGAACCGCCGCGCTGCTTTTGAACAGATGAAATACCC	4029
Oy	719	tgaataaagagagagatatgatgtcagttgttagatgtgcagagagatgttgat	778
Db	4030	TCATGAAAACGCTGACAGAGTGATTCGACCAATTCGACAGCATCTTCGATCGATT	4089
Oy	779	tgcttgggttaattccctgaagattctgaagttatcgaagacgaatcgaaggtttccgc	838
Db	4090	TGCTCGGAATCCTGGCTGATGATGAAGATCATTAAGCTTCCAATCATGCGCAACCA	4149
Oy	839	ttgtctctgaa	848
Db	4150	TTGCGATGGA	4159

Search completed: January 8, 2001, 09:20:44
Job time: 20658 sec

Run on: January 8, 2001, 03:42:11 ; Search time 178.73 seconds

Cell	Time	Updates/Sec
1	118.73	3094.400
2	118.73	3094.400
3	118.73	3094.400
4	118.73	3094.400
5	118.73	3094.400
6	118.73	3094.400
7	118.73	3094.400
8	118.73	3094.400
9	118.73	3094.400
10	118.73	3094.400
11	118.73	3094.400
12	118.73	3094.400
13	118.73	3094.400
14	118.73	3094.400
15	118.73	3094.400
16	118.73	3094.400
17	118.73	3094.400
18	118.73	3094.400
19	118.73	3094.400
20	118.73	3094.400
21	118.73	3094.400
22	118.73	3094.400
23	118.73	3094.400
24	118.73	3094.400
25	118.73	3094.400
26	118.73	3094.400
27	118.73	3094.400
28	118.73	3094.400
29	118.73	3094.400
30	118.73	3094.400
31	118.73	3094.400
32	118.73	3094.400
33	118.73	3094.400
34	118.73	3094.400
35	118.73	3094.400
36	118.73	3094.400
37	118.73	3094.400
38	118.73	3094.400
39	118.73	3094.400
40	118.73	3094.400
41	118.73	3094.400
42	118.73	3094.400
43	118.73	3094.400
44	118.73	3094.400
45	118.73	3094.400
46	118.73	3094.400
47	118.73	3094.400
48	118.73	3094.400
49	118.73	3094.400
50	118.73	3094.400
51	118.73	3094.400
52	118.73	3094.400
53	118.73	3094.400
54	118.73	3094.400
55	118.73	3094.400
56	118.73	3094.400
57	118.73	3094.400
58	118.73	3094.400
59	118.73	3094.400
60	118.73	3094.400
61	118.73	3094.400
62	118.73	3094.400
63	118.73	3094.400
64	118.73	3094.400
65	118.73	3094.400
66	118.73	3094.400
67	118.73	3094.400
68	118.73	3094.400
69	118.73	3094.400
70	118.73	3094.400
71	118.73	3094.400
72	118.73	3094.400
73	118.73	3094.400
74	118.73	3094.400
75	118.73	3094.400
76	118.73	3094.400
77	118.73	3094.400
78	118.73	3094.400
79	118.73	3094.400
80	118.73	3094.400
81	118.73	3094.400
82	118.73	3094.400
83	118.73	3094.400
84	118.73	3094.400
85	118.73	3094.400
86	118.73	3094.400
87	118.73	3094.400
88	118.73	3094.400
89	118.73	3094.400
90	118.73	3094.400
91	118.73	3094.400
92	118.73	3094.400
93	118.73	3094.400
94	118.73	3094.400
95	118.73	3094.400
96	118.73	3094.400
97	118.73	3094.400
98	118.73	3094.400
99	118.73	3094.400

```

Title:      US-09-553-431-1
Perfect score: 978
Sequence:   1 atgacgcctctgagattgtt.....tctctctctcttaccac 978

```

Scoring table:

	IDENTITY_NUC
Count	100

480022 seqs, 187831343 residues

number of milks satisfying chosen parameters: 960044

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database

1:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1980.DAT.*
2:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1981.DAT.*
3:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1982.DAT.*
4:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1983.DAT.*
5:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1984.DAT.*
6:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1985.DAT.*
7:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1986.DAT.*
8:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1987.DAT.*
9:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1988.DAT.*
10:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1989.DAT.*
11:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1990.DAT.*
12:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1991.DAT.*
13:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1992.DAT.*
14:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1993.DAT.*
15:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1994.DAT.*
16:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1995.DAT.*
17:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1996.DAT.*
18:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1997.DAT.*
19:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1998.DAT.*
20:	/cgm2_2/gcgdata/gemseq/gemseqn/NA1999.DAT.*
21:	/cgm2_2/gcgdata/gemseq/gemseqn/NA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result							
No.	Score	Query Match	Length	DB	ID	Description	
1	42.8	4.4	885	19	X30457	H. pylori cellular	
2	42.8	4.4	891	18	T67691	H. pylori cytoplasm	
3	42.4	4.3	513	18	T67662	H. pylori cytoplasm	
4	42.4	4.3	6761	20	X20517	Polynucleotide seq	
5	42.2	4.3	1236	18	T67616	H. pylori cytoplasm	
6	42.2	4.3	1275	18	T68266	H. pylori cytoplasm	
7	39.2	4.0	780	20	V90595	Nucleotide sequence	
8	39	4.0	1107	19	V38816	DNA encoding a dur	
9	38.4	3.9	1664	976	19	V21209	Methanococcus jann
10	38.2	3.9	1012	20	V90875	Nucleotide sequence	
11	37.4	3.8	30600	20	Z32023	Human MERT1 relate	
12	35.6	3.6	6659	21	Z36862	DNA encoding a GON	

13	3.5	3.6	1590	20	X07125	Staphylococcus aur
14	35	3.6	7778	18	V74359	Staphylococcus aur
15	34.6	3.5	534720	19	V30458	Rhizobium species
16	34.6	3.5	536165	19	V30459	Rhizobium species
17	33.6	3.4	44377	18	T78508	Platenolide synth
18	33.6	3.4	44377	18	T80414	Platenolide synth
19	33.4	3.4	1610	18	X13498	Enterococcus faeca
20	33	3.4	58407	19	V21210	Enterococcus faeca
21	32.8	3.4	4652	21	Z42709	Human 5' EST
22	32.8	3.4	1128	19	Z42709	Human 5' EST
23	32.8	3.4	10401	18	X13211	Staphylococcus aur
24	32.8	3.4	12022	20	V74460	Enterococcus faeca
25	32.2	3.3	1866	21	X13054	Human protein Kina
26	32.2	3.3	26281	21	Z68995	Nucleotide sequenc
27	32	3.3	15062	19	V44332	Lettuce resistance
28	32	3.3	116277	20	X20249	Borrelia burgdorfe
29	31.8	3.3	7400	18	V77264	Staphylococcus aur
30	31.8	3.3	7032	20	V99601	Canine factor VIII
31	31.8	3.3	7032	21	Z56579	Canine factor VIII
32	31.8	3.3	7494	18	V74380	Staphylococcus aur
33	31.8	3.3	9223	19	V52192	Staphylococcus aur
34	31.8	3.3	1664976	19	V21209	Staphylococcus pneu
35	31.4	3.2	1109	18	V74533	Methanococcus jann
36	31.4	3.2	72928	20	V74533	Staphylococcus aur
37	31.2	3.2	1967	20	V64414	Human ASTHJ 5' ge
38	31.2	3.2	26016	19	V57272	Human PTRAJ-1 gene
39	31	3.2	1690	18	V74797	Human flavin-con
40	31	3.2	9897	19	V52143	Staphylococcus aur
41	30.8	3.1	1098	18	T67896	Staphylococcus pneu
42	30.8	3.1	3786	19	V69424	H. pylori cytoplas
43	30.8	3.1	3786	21	Z89989	Escherichia sp. exte
44	30.8	3.1	7664	20	X13512	Escherichia antigen
45	30.6	3.1	5678	20	X13570	Enterococcus faeca

ALIGNMENTS

RESULT	1
ID	X30457 standard; DNA; 885 BP.
XX	
AC	X30457;
XX	
DT	08-JUN-1999 (first entry)
XX	
DE	
XX	
XX	H. pylori cellular protein ORF 06ep11202_4569693_c2_28.
XX	
KW	Vaccine; probe; diagnostic; ORF; cell envelope protein;
KW	Secreted protein; cellular protein; ds.
XX	
XX	Helicobacter pylori.
OS	
XX	
PN	W09818323-A1.
XX	
PD	07-MAY-1998.
XX	
PF	28-OCT-1997; 97MO-US19575.
XX	
PR	14-JUL-1997; 97US-0891928.
PR	28-OCT-1996; 96US-0739150.
PR	06-DEC-1996; 96US-0759739.
XX	
PA	(ASTR) ASTRA AB.
XX	
PI	Aim RA, Smith D;
XX	
DR	WPI; 1998-271811/24.
DR	P-PSDB; Y10990.
XX	
PT	Helicobacter pylori nucleic acids and proteins - used to develop
PT	products for the detection, prevention and treatment of H. pylori
PT	infections

XX

XX
PN
XX
WO9640893-A1

XX

XX
PN
XX
WO9640893-A1

XX (ASTR) ASTRA AB.
PA Berglindh OT, Smith D, Mellgaerd BL.
XX WPI: 1997-052306/05.
DR P-PSDB; W20443.
XX
XX Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
XX
XX Claim 9; Pages 267; 1481pp; English.
XX
XX The present sequence encodes a Helicobacter pylori cytoplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds, useful
CC as potential H. pylori life cycle activators or inhibitors. The genomic
CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF of at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
XX
SQ Sequence 1236 BP; 398 A; 239 C; 282 G; 317 T; 0 other;

Query Match
Best Local Similarity 4.3%; Score 42.2; DB 18; Length 1236;
Matches 92; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 155 aactgcgcgagaaacgcgcatcgtatcgtatcacctccggaagaagcggttgaa 214
D 404 atctgctaataaacatcaacatgctgcatgataagccaagcggttgaa 463
QY 215 agacgaacaccccgcaaatgctcgtctctctcgtcgttcggttcggttcg 274
D 464 aaagacacaccccgcaaatgctcgtctctcgtcgttcggttcggttcg 523
QY 275 ccattgacgcgcgacgttgcgtcgtacacgtcctcctcgaaggttagaa 329
D 524 tactagaacgctgattgctatgagccctaatatccctagatgaggttcgaa 578

RESULT 6
T68266 ID T68266 standard; DNA; 1275 BP.
XX
AC T68266;
XX
D 22-JUL-1997 (first entry)
XX
DE H. pylori cytoplasmic protein ORF hpse1521lorf13.
XX
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; ds.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT 1..1275
FT CDS
XX /*tag= a
XX /note= "no stop codon given in sequence"
XX
PN M09640893-A1.
XX
PD 19-DEC-1996.

XX
PF 06-JUN-1996; 96WO-US09122.
XX
PR 01-APR-1996; 96US-0630405.
PR 07-JUN-1995; 95US-0487032.
XX
PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL.
XX WPI: 1997-052306/05.
DR P-PSDB; W21013.
XX
XX Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
XX
XX Claim 9; Pages 1001; 1481pp; English.
XX
XX The present sequence encodes a Helicobacter pylori cytoplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds, useful
CC as potential H. pylori life cycle activators or inhibitors. The genomic
CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF of at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
XX
SQ Sequence 1275 BP; 410 A; 246 C; 289 G; 330 T; 0 other;

Query Match
Best Local Similarity 4.3%; Score 42.2; DB 18; Length 1275;
Matches 92; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 155 aactgcgcgagaaacgcgcatcgtatcgtatcacctccggaagaagcggttgaa 214
D 443 atctgctaataaacatcaacatgctgcatgataagccaagcggttgaa 502
QY 215 agacgaacaccccgcaaatgctcgtctctcgtcgttcggttcggttcg 274
D 503 aaagacacaccccgcaaatgctcgtctctcgtcgttcggttcggttcg 562
QY 275 ccattgacgcgcgacgttgcgtcgtacacgtcctcctcgaaggttagaa 329
D 563 tactagaacgctgattgctatgagccctaatatccctagatgaggttcgaa 617

RESULT 7
V90595/C ID V90595 standard; DNA; 780 BP.
XX
AC V90595;
XX
D 18-FEB-1999 (first entry)
XX
DE Nucleotide sequence of clone Y108-1.ASM from cluster 32.
XX
KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma; ss.
XX
OS Helicobacter pylori.
XX
PN M09849314-A2.
XX
PD 05-NOV-1998.

27-APR-1998; 98WO-US08487.
 14-OCT-1997; 97US-0061958.
 25-APR-1997; 97US-0045107.
 (GENE-) GENELABS TECHNOLOGIES INC.

Chow TP, Fry KE, Lim MY, McAttee CP;
 WPI; 1999-009433/01.

New *Helicobacter pylori* antigens and related nucleic acid sequences
 - useful in serological diagnosis and protective vaccines, providing
 long-lasting immune response
 Claim 27; Page 125; 402pp; English.

The present sequence encodes a *Helicobacter pylori* antigenic protein
 that is characterised by immunoreactivity with *H. pylori*-positive
 antisera. The proteins are highly immunogenic and induce a long-lasting
 immune response that persists even after antimicrobial treatment. In
 antibody-detection assays, on sera, plasma, urine, saliva etc., they are
 highly sensitive and specific. The specification also describes 69
 previously unrecognised immunogenic cluster families. *H. pylori* antigens
 are used to detect *H. pylori*-specific infection, for diagnosing
 infection or to confirm eradication of infection, and in vaccines to
 protect against *H. pylori* infection and related diseases (gastritis,
 peptic ulcer, gastric adenocarcinoma/lymphoma).

Sequence 780 BP; 218 A; 157 C; 142 G; 263 T; 0 other;

Query Match
 Best Local Similarity 4.0%; Score 39.2; DB 20; Length 780;
 Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 174 gcatatcgctgctacacccggaagagcggtgtggaagagagacacacggcaaa 233
 DB 296 GCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
 QY 234 tctcgcgtcctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 293
 DB 236 TTTTAAAGCATGCTTTTAAAGCATGCTTTTAAAGCATGCTTTTAAAGCATGCT 237
 QY 294 tctcgcgtcctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 329
 DB 176 TGGCCCTAATATATCCCTAGATGATGATGATGATGATGATGATGATGATGAT 141

RESULT 8
 V38816/C
 ID V38816 standard; DNA; 1107 BP.
 AC V38816;
 XX
 DT 09-OCT-1998 (first entry)
 XX
 DE DNA encoding a durum wheat glutenin protein.
 XX
 KW Glutenin gene: durum wheat; low-molecular-weight;
 KW transgenic durum wheat; ss.
 XX
 OS Triticum durum.
 XX
 PN FR2757538-A1.
 XX
 PD 26-JUN-1998.
 XX
 PF 18-DEC-1997; 97FR-0016059.
 XX
 PR 19-DEC-1996; 96IT-MI02663.
 XX
 PA (TUVY-) ITAL MIN UNIV RICERCA SCI & TECNOLOGICA.

Cardelli LE, D'Ovidio R, Marchitelli C, Porceddu E;
 WPI; 1998-365055/32.
 P-PSDB; W62647.

Durum wheat glutenin gene - coding for glutenin protein of low
 molecular weight

Claim 1; Page 13; 18pp; French.

The present sequence represents a glutenin gene, and is isolated from
 the genomic DNA of Triticum durum L. The gene codes for a
 low-molecular-weight glutenin protein and can be used to produce
 transgenic durum wheat plants with "better quality characteristics"
 [no details given].

Sequence 1107 BP; 368 A; 377 C; 152 G; 210 T; 0 other;

Query Match
 Best Local Similarity 4.8%; Score 39; DB 19; Length 1107;
 Matches 150; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 600 agtctcgttaacacactccgagatatacagcgttaagagatgctgataagggttacgggtt 659
 DB 380 AATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 321
 QY 660 gctcgtatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 719
 DB 320 GGTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
 QY 720 gatlaagagagagatagatagatagatagatagatagatagatagatagatagat 779
 DB 260 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
 QY 780 gctcgtatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 839
 DB 200 GGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 141
 QY 840 tctcgtatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 899
 DB 140 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 81
 QY 900 cgttcgagcaagatagatagatagatagatagatagatagatagatagatagat 934
 DB 80 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 46

RESULT 9
 V21209/C
 ID V21209 standard; DNA; 1664976 BP.
 AC V21209;
 XX
 DT 10-NOV-1998 (first entry)
 XX
 DE Methanococcus jannaschii circular chromosome.
 XX
 KW Methanococcus jannaschii; methanogenic archaeon; circular chromosome;
 KW genome; autotrophic; extrachromosomal element; identification; ds.
 XX
 OS Methanococcus jannaschii.
 XX
 PN W09807830-A2.
 XX
 PD 26-FEB-1998.
 XX
 PF 22-AUG-1997; 97WO-US14900.
 XX
 PR 22-AUG-1996; 96US-0024428.
 XX
 PA (GENO-) INST GENOMIC RES.

PA (UNIT) UNIV ILLINOIS FOUND.
 PA (UYJO) UNIV JOHN HOPKINS SCHOOL MEDICINE.
 XX
 XX Bult CJ, Smith HO, Venter JC, White OR, Woese CR.
 PI WPI, 1998-169145/15.
 XX
 DR Complete genome sequence of methano-genic archaeon, Methanococcus
 PT jannaschii - useful in identification of M. jannaschii genome
 PT fragment
 XX
 PS Claim 13; Page 152-585; 614pp; English.
 XX
 CC The present sequence represents the complete 1.66-megabase pair genome
 CC sequence of the Methanococcus jannaschii open reading frames from the
 CC present invention describes M. jannaschii a computer based system
 CC genome sequence. The invention also describes a computer based system
 CC for identifying fragments of the M. jannaschii genome that are
 CC homologous to target nucleotide sequences, comprising: (a) data storage
 CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550
 CC bp sequence (see V21209, V21210 and V21211), or a nucleotide sequence at
 CC least 99.9% identical to it; (b) search means for comparing a target
 CC sequence to the nucleotide sequence of the data storage means to
 CC identify a homologous sequence, and (c) retrieval means for obtaining
 CC the homologous sequence. The method, which is based on whole genome
 CC random sequencing of an autotrophic archaeon M. jannaschii, the genome
 CC of which consists of 3 physically distinct elements, a large circular
 CC chromosome (the 1664976 bp sequence given in V21209), a large circular
 CC extra-chromosomal element (the 58407 bp sequence given in V21210), and a
 CC small circular extra-chromosomal element (the 16550 bp sequence given in
 CC V21211), can be used in the identification of M. jannaschii genome
 CC fragment.
 XX
 SQ Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other:
 XX
 Query Match 3.9%; Score 38.4; DB 19; Length 1664976;
 Best Local Similarity 54.2%; Pred. No. 1.5;
 Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 186 tateactccggaagaagcggtgtgtggaagacgacacaccgcaatgctcctc 245
 DB 1301097 TATAGCTTCGGGTAAAGAGCGGTGGAAACATCAATCAGCATCTTTAGCGAGTAC 1301038
 QY 246 tctcgctgcttaacggttctcgtctgcacatgacgcgcacgtgctccgaact 305
 DB 1301037 ACTTGTCTAAATTTGGGAAAAAGGTTTGTAGCTATTGATGAGACATATCATGCGTAATTT 1300978
 QY 306 cgatctcctcctagaggttagaga 329
 DB 1300977 AGGGAATTCCTATTCATATGAAAA 1300954
 XX
 RESULT 10
 V90875/c
 ID V90875 standard; DNA; 1012 BP.
 XX
 AC V90875;
 XX
 DT 18-FEB-1999 (first entry)
 XX
 DE Nucleotide sequence of cluster 32.
 XX
 KW Antigen: immunogenic cluster family; vaccine; gastritis; diagnosis;
 KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma; ss.
 XX
 OS Helicobacter pylori.
 XX
 OS WO9849314-A2.
 XX
 PN 05-NOV-1998.
 XX
 PD 27-APR-1998; 98WO-US08487.
 XX
 PF

XX
 PR 14-OCT-1997; 97US-0061958.
 PR 25-APR-1997; 97US-0045107.
 XX
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 XX
 PI Chow TP, Fry KE, Lim MY, McAttee CP;
 XX
 DR WPI, 1999-009433/01.
 XX
 PT New Helicobacter pylori antigens and related nucleic acid sequences
 PT - useful in serological diagnosis and protective vaccines, providing
 PT long-lasting immune response
 XX
 PS Claim 23; Page 297; 402pp; English.
 XX
 CC The present sequence encodes a Helicobacter pylori antigenic protein
 CC that is characterised by immunoreactivity with H. pylori-positive
 CC antisera. The proteins are highly immunogenic and induce a long-lasting
 CC immune response that persists even after antimicrobial treatment. In
 CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
 CC highly sensitive and specific. The specification also describes 69
 CC previously unrecognised immunogenic cluster families. H. pylori antigens
 CC are used to detect H. pylori-specific antibodies, for diagnosing
 CC infection or to confirm eradication of infection, and in vaccines to
 CC protect against H. pylori infection and related diseases (gastritis,
 CC peptic ulcer, gastric adenocarcinoma/lymphoma).
 XX
 SQ Sequence 1012 BP; 293 A; 188 C; 197 G; 333 T; 1 other:
 XX
 Query Match 3.9%; Score 38.2; DB 20; Length 1012;
 Best Local Similarity 52.9%; Pred. No. 0.033;
 Matches 82; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 QY 174 gctatcgtctatcacccctccggaagaagcggtgttggaaagacgacaccgcaaa 233
 DB 304 GCATGTAGTATGATATGATAGCTCAGGTAAAGCGGTGGTAAAGCACACCGCTGAA 245
 QY 234 tgcgtctcctcctcgtcgttgcaggttctcagttgtgcacatgacgcgcacttgg 293
 DB 244 TTTAAGCATCGCTTTAGCGAATTTAAACCAAAAGTGGGTTACTTACGACGCTGATGTGA 185
 QY 294 tctcgaactcgtatcctcccttaggttagaga 328
 DB 184 TGGCCCTAATATCCCTAGATGATGGCTTGCAAA 150
 XX
 RESULT 11
 Z32023/c
 ID Z32023 standard; DNA; 30600 BP.
 XX
 AC Z32023;
 XX
 DT 10-JAN-2000 (first entry)
 XX
 DE Human METH1 related EST Z69361.
 XX
 KW Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
 KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
 KW angiogenesis inhibitor; abnormal wound healing; inflammation;
 KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
 KW diabetic retinopathy; macula degeneration; haemangioma; detection;
 KW arterial-venous malformation; immune deficiency; ss.
 XX
 OS Homo sapiens.
 XX
 OS WO9937660-A1.
 XX
 PN 29-JUL-1999.
 XX
 PD 22-JAN-1999; 99WO-US01313.
 XX
 PF

PR 23-JAN-1998: 98US-0072298.
 XX 28-AUG-1998: 98US-0098539.
 PA (TRUE/) IRUELA-ARISPE L.
 PA (HAST/) HASTINGS G A.
 PA (RUBE/) RUBEN S M.
 PI Iruele-Arispe L, Hastings GA, Ruben SM;
 DR WPI: 1999-590684/50.
 XX
 PT New isolated metalloprotease thrombospondin polypeptides, useful for
 PT treating hyperproliferative disorders, cancers or autoimmune disorders
 PS
 XX Disclosure: Page 327-347; 457pp; English.
 CC 232000 and 232001 encode, and Y49501 and Y49502 represent, human
 CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
 CC respectively. METH1 and METH2 have been found to be potent inhibitors of
 CC angiogenesis both in vitro and in vivo. They can be used for treating
 CC cancer and other disorders related to angiogenesis including abnormal
 CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
 CC endometrial bleeding disorders, diabetic retinopathy, some forms of
 CC macular degeneration, haemangiomas, and arterial-venous malformations.
 CC They may be useful in treating deficiencies or disorders of the immune
 CC system, by activating or inhibiting the proliferation, differentiation,
 CC or mobilisation (chemotaxis) of immune cells. The etiology of these
 CC immune deficiencies or disorders may be genetic, somatic, such as
 CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
 CC toxins), or infectious. They can also be used to treat inflammatory or
 CC conditions, both chronic and acute conditions. The products can also be
 CC used for detection and diagnosis. 232002 to 232080, and Y49503 to Y49511
 CC represent sequences given in the exemplification of the present
 CC invention.
 XX
 SG Sequence 30600 BP; 10390 A; 5313 C; 5429 G; 9467 T; 1 other:
 Query Match 3.8%; Score 37.4; DB 20; Length 30600;
 Best Local Similarity 50.2%; Pred. No. 0.37;
 Matches 118; Conservative 0; Mismatches 116; Indels 1; Gaps 1;
 QY 322 ttgaggaatcggatcattcacttcgtcgtgaggtataaaggagatgctgcctgat 381
 DB 1794 ttgcataaacatctcattcagcattcctcaataatataaacttggagaaatggtctcact 1735
 QY 382 caagctctggtacgtatgaataagcgttggcgaattcgaattgctatatacctaaccct 441
 DB 1734 gtatg-tgtttcttggatgagatggtcctcaagaagatttcattggttctggaaat 1676
 QY 442 agatcgaaacttcgagttgagttgtgtggaagcattggaattgagttgtgagtcgttg 501
 DB 1675 ggaaactgaaagtcgattgtcactttgtgtggtatcattctgattcggcctgcttcacag 1616
 QY 502 aaactcgaacggaggttccacggattcattcattcattcattcattcattcattcattc 556
 DB 1615 acattgtaattttagcacttgcctcatttttgaggaataatgattgagactgctgtaa 1561
 RESULT 12
 ID 236862 standard; DNA: 6659 BP.
 XX 236862:
 AC
 XX 13-MAR-2000 (first entry)
 DT
 XX
 DE DNA encoding a GON-1 protein of *Caenorhabditis elegans*.
 XX
 KW GON-1; metalloprotease; cell migration; modulator;
 KW metalloprotease domain; thrombospondin domain; abnormal cell migration;
 KW organ shaping; sterility; cancer metastasis; ss.

XX
 OS *Caenorhabditis elegans*.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..6453
 FT /*tag= a
 FT /product= "GON-1"
 XX
 PN W0961656-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 28-MAY-1999: 99NO-US11918.
 XX
 PR 29-MAY-1998: 98US-0087170.
 PR 13-APR-1999: 99US-0129023.
 XX
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 PI Kimble JE, Bieleloch RH;
 DR WPI: 2000-072633/06.
 DR E-PSDB: Y53898.
 PT Identifying modulators of proteins containing metalloprotease and
 PT thrombospondin domains, potentially useful for controlling cell
 PT migration and organ shaping
 PS
 XX Disclosure: Page 39-51; 60pp; English.
 CC The present sequence encodes a GON-1 protein of *Caenorhabditis elegans*.
 CC GON-1 is a secreted metalloprotease that lacks a transmembrane domain
 CC and possesses a predicted metalloprotease domain between amino acids
 CC 269-456. In *C. elegans* hermaphrodites, GON-1 is required for migration of
 CC two distal tip cells to produce elongated tubes, whereas in males, GON-1
 CC is required for migration of a single linker cell to produce a single
 CC elongated tube. The protein is used in the method of the invention.
 CC specification describes a method for identifying a modulator of a
 CC protein that contains a metalloprotease domain and a thrombospondin
 CC domain. The method comprises treating a target organism, having a
 CC developing gonadal cell that is responsive to the protein, with a test
 CC compound, and determining any change in migration or shape of the cell
 CC attributable to the test compound. The compounds identified are
 CC potential therapeutic modulators of abnormal cell migration and organ
 CC shaping, e.g. for rendering animals (specifically nematodes) sterile
 CC and for inhibiting cancer metastases.
 XX
 SG Sequence 6659 BP; 1928 A; 1337 C; 1547 G; 1847 T; 0 other:
 Query Match 3.6%; Score 35.6; DB 21; Length 6659;
 Best Local Similarity 50.0%; Pred. No. 0.61;
 Matches 115; Conservative 0; Mismatches 114; Indels 1; Gaps 1;
 QY 322 ttgaggaatcggatcattcacttcgtcgtgaggtataaaggagatgctgcctgat 381
 DB 3479 ttacaaaccatgctcgaataagaaatcgtccaaataaacttggagaaatggtctcagt 3538
 QY 382 caagctctggtacgtatgaataagcgttggcgaattcgaattgctatatacctaaccct 441
 DB 3539 gtatg-tgttcttggatgagatggtcgtcgaagaagtttcattggttctggaaat 3597
 QY 442 agatcgaaacttcgagttgagttgtgtggaagcattggaattgagttgtgagtcgttg 501
 DB 3598 ggaactggaagtcgatactgctcatttgggttactgcatcgcgtcgtccttcacacag 3657
 QY 502 aaactcgaacggaggttccacggattcattcattcattcattcattcattcattcattc 551
 DB 3658 acattgtaatttaggacttgcctcatttttgaggaataatcattcattcattcattcattc 3707
 RESULT 13
 X07125

ID X07125 standard; DNA; 1590 BP.
 XX X07125;
 AC
 XX
 DT 21-MAY-1999 (first entry)
 XX
 DE Staphylococcus aureus mutant P11C12 virulence gene.
 XX
 KM Virulence; P11C12; vaccine; antibacterial; antibiotic;
 KM screening; bacteremia; ss.
 XX
 OS Staphylococcus aureus.
 XX
 PN W09901473-A2.
 XX
 PD 14-JAN-1999.
 XX
 PF 03-JUL-1998; 98WO-GB01974.
 XX
 PR 03-JUL-1997; 97US-0887534.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Holden DW;
 XX
 DR WPI: 1999-105999/09.
 XX
 DT P-PSDB; W97715.
 XX
 PT Inhibition of virulence genes from Staphylococcus aureus - useful
 PT for, e.g. screening for potential anti-microbial agents
 PS
 XX Claim 1; Page 152-154; 203pp; English.
 CC This is the nucleotide sequence of a virulence gene identified in
 CC Staphylococcus aureus mutant P11C12. S. aureus genes (see X07088-136)
 CC associated with virulence were identified by signature-tagged
 CC mutagenesis in which mutants containing a chromosomal insertion of
 CC a signature tagged transposon were generated, mutants with
 CC attenuated virulence were identified in a mouse model of bacteraemia,
 CC and the nucleotide sequences of the regions flanking the transposon
 CC insertion sites of these mutants were determined. Database sequence
 CC comparisons were performed to identify the virulence genes and to
 CC determine the possible function of their protein products (see also
 CC W97680-724). The P11C12 virulence gene product (see W97715) is
 CC unknown, but has 40% identity to Mycoplasma mycoides hypothetical
 CC protein in ffn 5' region q0144. A claimed method of identifying
 CC an antibacterial agent involves assaying potential agents for the
 CC ability to interfere with the expression of S. aureus virulence
 CC gene products. Also new is S. aureus containing a functional
 CC mutation in one of the virulence genes, and its use in vaccine
 CC compositions.
 XX
 SO Sequence 1590 BP; 617 A; 232 C; 340 G; 401 T; 0 other;

Query Match 3.6%; Score 35; DB 20; Length 1590;
 Best Local Similarity 48.3%; Pred. No. 0.45;
 Matches 98; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 595 gaagcagttctgtgaacaccccgagataacagcggttaaggaagtcgtagggttacg 654
 DB 778 gaagcggttatattatctggtttgatacaataagaagaagaatgctagaacagcactt 837
 QY 655 gtttggtaaatgcatggaatcagagataataagaatgagtggaacagtgagaa 714
 DB 838 gttactatgatactgtagcagcattatccaggttaaatggaatgataatggtcga 897
 QY 715 gatatgataaaggagagatagatgacgtgttagatgtgacagagaaatgttgagtg 774
 DB 898 gctagaagaagagtagacgatatattagagaagagtggaacagcctacattgaagtg 957
 QY 775 tcattgcttggtgaattccctga 797

DB 958 aacgcataatagcatcctga 980

RESULT 14
 ID V74359 standard; DNA; 7778 BP.
 XX V74359;
 AC
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus contig SEQ ID #48.
 XX
 DE Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.
 XX
 XX Key Location/Qualifiers
 FH 1681..1740
 FT misc-feature
 FT /tag= a
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc-feature
 FT 3481..3540
 FT /tag= b
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc-feature
 FT 5281..5340
 FT /tag= c
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc-feature
 FT 7081..7140
 FT /tag= d
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"

EP786519-A2.
 30-JUL-1997.
 PF 07-JAN-1997; 97EP-0100117.
 PR 05-JAN-1996; 96US-0009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 XX WPI: 1997-374922/35.
 DR
 XX
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines
 XX
 PS Claim 1; Page 387-391; 3271pp; English.
 XX
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S. aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or


```
FT /gene= "fixA"  
FT /product= "protein required for nitrogenase activity"  
FT complement (442316..442636)  
FT /tag= t  
FT /standard_name= "ORF K20"  
FT /product= "protein of unknown function"  
FT complement (443313..443879)  
FT /tag= u  
FT /standard_name= "ORF K21"  
FT /product= "protein of unknown function"  
FT 444337..445029  
FT /tag= v  
FT /standard_name= "ORF K22"  
FT /product= "ferrodoxin-like protein"  
FT /note= "homologous to the NifH gene"  
FT 445088..446602  
FT /tag= w  
FT /standard_name= "ORF K23"  
FT /gene= "dctA"  
FT /product= "C4-dicarboxylate transport protein"  
FT /note= "homologous to the DctAI gene"  
FT 446599..447843  
FT /tag= x  
FT /standard_name= "ORF L1"  
FT /product= "cytochrome P450-like protein"  
FT /note= "homologous to the CamC gene"  
FT 447844..448500  
FT /tag= y  
FT /standard_name= "ORF L2"  
FT /product= "gamma-hexachlorocyclohexane-dechlorinase-like  
FT protein"  
FT /note= "homologous to the LinA gene"  
FT 448497..450203  
FT /tag= z  
FT /standard_name= "ORF L3"  
FT /product= "putative protein with degradative function"  
FT 450341..451396  
FT /tag= aa  
FT /standard_name= "ORF L4"  
FT /product= "luciferase alpha-subunit-like protein"  
FT /note= "homologous to the luxA gene"  
FT 452980..454494  
FT /tag= ab  
FT /standard_name= "ORF L6"  
FT /gene= "nifD"  
FT /product= "alpha-subunit of fmo protein of nitrogenase"  
FT 454590..456131  
FT /tag= ac  
FT /standard_name= "ORF L7"  
FT /gene= "nifK"  
FT /product= "beta-subunit of fmo protein of nitrogenase"  
FT 456187..457677  
FT /tag= ad  
FT /standard_name= "ORF L8"  
FT /product= "protein involved in fmo co-factor  
FT biosynthesis"  
FT /note= "homologous to the NifB gene"  
FT 457687..459096  
FT /tag= ae  
FT /standard_name= "ORF L9"  
FT /product= "protein involved in fmo co-factor  
FT biosynthesis"  
FT /note= "homologous to the FixF gene"  
FT 459093..459575  
FT /tag= af  
FT /standard_name= "ORF L10"  
FT /product= "protein of unknown function"  
FT /note= "homologous to the NifX gene"  
FT 459579..460067  
FT /tag= ag  
FT /standard_name= "ORF L11"  
FT /product= "protein of unknown function"  
FT 460501..460920
```

```
FT /tag= ah  
FT /standard_name= "ORF L12"  
FT /product= "protein similar to part of the fe protein  
FT of nitrogenase"  
FT 461228..461545  
FT /tag= ai  
FT /standard_name= "ORF L13"  
FT /product= "protein of unknown function"  
FT 461201..464739  
FT /tag= aj  
FT /standard_name= "ORF L14"  
FT /product= "peptidase-like protein"  
FT /note= "homologous to the bi-MPP gene"  
FT 464736..466079  
FT /tag= ak  
FT /standard_name= "ORF L15"  
FT /product= "processing protease-like protein"  
FT /note= "homologous to the pp gene"  
FT 466590..467021  
FT CDS  
  
Query Match 3.5%; Score 34.6; DB 19; Length 534720;  
Best Local Similarity 54.3%; Pred. No. 13;  
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
  
QY 161 ccgagaaacgcgcgtatcgttcaccccgaaagcggtgtgaaagcga 220  
DB 56482 CGGGGANAAGCTCCAGTCATGCCGTTGTCATTTTCAAGCGGTTCCGANAAGCA 56423  
  
QY 221 caaccacgcgaagtcgtctctctcgcgtcgttaaggttccagttgcgcattg 280  
DB 56422 CGAGCGCCGCCATCTCCGCGACATACATGCGGTGACCGGACCGGTCTTGCGCGTG 56363  
  
QY 281 acgcgcgacc 289  
DB 56362 ACCTCGACC 56354
```

Search completed: January 8, 2001, 09:29:10
Job time: 20819 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2001, 03:38:45 ; Search time 81.49 Seconds
(without alignments)
1934.161 Million cell updates/sec

Title: US-09-553-431-1

Perfect score: 978
Sequence: 1 atggcgctctctgagatgttcttctctctcttggcggc 978

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2-6/ptodata/2/lna/5A_COMB.seq.*
2: /cgn2-6/ptodata/2/lna/5B_COMB.seq.*
3: /cgn2-6/ptodata/2/lna/6_COMB.seq.*
4: /cgn2-6/ptodata/2/lna/PCRUS_COMB.seq.*
5: /cgn2-6/ptodata/2/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
C 1	46.6	4.8	7218	1	US-08-232-463-14 Sequence 14, Appl
C 2	39	4.0	1107	2	US-08-991-300-1 Sequence 1, Appl
C 3	34	3.5	7218	1	US-08-232-463-14 Sequence 14, Appl
C 4	33.6	3.4	44377	2	US-08-804-227C-7 Sequence 7, Appl
C 5	33.6	3.4	44377	2	US-08-804-227C-7 Sequence 7, Appl
C 6	32.2	3.3	1866	3	US-09-173-581-13 Sequence 13, Appl
C 7	31.4	3.2	72928	3	US-09-009-913-1 Sequence 1, Appl
C 8	31.2	3.2	1964	2	US-08-841-343-8 Sequence 8, Appl
C 9	30.2	3.1	1041	3	US-09-033-055A-4 Sequence 4, Appl
C 10	29.6	3.0	4928	1	US-08-399-561-1 Sequence 1, Appl
C 11	29.6	3.0	5000	1	US-09-147-522-5 Sequence 5, Appl
C 12	29.6	3.0	246240	2	US-08-724-394A-20 Sequence 20, Appl
C 13	29.6	3.0	246240	2	US-08-724-394A-21 Sequence 21, Appl
C 14	29.6	3.0	246240	2	US-08-724-394A-22 Sequence 22, Appl
C 15	29	3.0	2325	2	US-08-966-388-5 Sequence 5, Appl
C 16	29	3.0	2325	2	US-08-966-388-5 Sequence 5, Appl
C 17	29	3.0	2325	3	US-09-188-403-5 Sequence 5, Appl
C 18	29	3.0	2325	3	US-09-188-403-5 Sequence 5, Appl
C 19	29	3.0	3956	2	US-08-966-388-9 Sequence 9, Appl
C 20	29	3.0	3956	2	US-08-966-388-9 Sequence 9, Appl
C 21	29	3.0	3956	3	US-09-188-403-9 Sequence 9, Appl
C 22	28.8	2.9	740	3	US-08-771-098-1 Sequence 1, Appl
C 23	28.8	2.9	740	3	US-08-771-098-1 Sequence 1, Appl
C 24	28.8	2.9	3126	3	US-09-022-875-3 Sequence 3, Appl
C 25	28.6	2.9	433	1	US-08-929-329-2 Sequence 2, Appl
C 26	28.6	2.9	613	1	US-08-428-733A-1 Sequence 1, Appl
C 27	28.6	2.9	2989	3	US-08-728-956-2 Sequence 2, Appl
C 28	28.6	2.9	3240	1	US-09-068-195-10 Sequence 10, Appl
C 29	28.6	2.9	3240	1	US-08-162-081B-34 Sequence 34, Appl
C 30	28.6	2.9	3240	2	US-08-780-872-34 Sequence 34, Appl

C 29	28.6	2.9	3412	1	US-08-162-081B-32 Sequence 32, Appl
C 30	28.6	2.9	3412	2	US-08-780-872-32 Sequence 32, Appl
C 31	28.6	2.9	5565	3	US-09-068-195-21 Sequence 21, Appl
C 32	28.6	2.9	9468	1	US-08-325-547-10 Sequence 10, Appl
C 33	28.6	2.9	13987	2	US-08-804-227C-13 Sequence 13, Appl
C 34	28.6	2.9	43280	2	US-08-804-227C-13 Sequence 13, Appl
C 35	28.6	2.9	1215	4	PCT-US96-05320A-707 Sequence 707, App
C 36	28.4	2.9	2101	3	US-08-637-763B-5 Sequence 5, Appl
C 37	28.4	2.9	2101	3	US-09-170-354-5 Sequence 5, Appl
C 38	28.4	2.9	9045	3	US-09-121-321-1 Patent No. 5171685
C 39	28.2	2.9	928	5	US-08-624-125-3 Patent No. 5171685
C 40	28.2	2.9	928	5	US-08-624-125-3 Patent No. 5171685
C 41	28.2	2.9	956	1	US-08-470-260-7 Sequence 7, Appl
C 42	28.2	2.9	1838	3	US-08-471-491-7 Sequence 7, Appl
C 43	28.2	2.9	1838	3	US-08-466-662-7 Sequence 7, Appl
C 44	28.2	2.9	1838	3	US-08-466-662-7 Sequence 7, Appl
C 45	28.2	2.9	3306	1	US-08-261-206A-71 Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT9gpt-Fls
US-08-232-463-14

Query Match

4.8% ; Score 46.6 ; DB 1 ; Length 7218 ;

us-09-553-431-1.rni

```

1  RESULT 5
2  US-08-804-198-1
3  : Sequence 1, Application US/080804198
4  : Patent No. 5945370
5  :
6  : GENERAL INFORMATION:
7  :
8  : APPLICANT: Burgett, Stanley G.
9  : APPLICANT: Knustoss, Stuart A.
10 : APPLICANT: Rao, Nagaraja R.
11 : APPLICANT: Richardson, Mark A.
12 : APPLICANT: Rostek, Paul R., Jr.
13 : TITLE OF INVENTION: PLATENOIDE SYSTE
14 :
15 : NUMBER OF SEQUENCES: 6
16 :
17 : CORRESPONDENCE ADDRESS:
18 :
19 : ADDRESSEE: PAUL R. CANTRELL, 1138
20 : STREET: LILLY CORPORATE CENTER
21 : CITY: INDIANAPOLIS
22 : STATE: IN
23 :
24 : COUNTRY: USA
25 : ZIP: 46285
26 :
27 : COMPUTER READABLE FORM:
28 : MEDIUM TYPE: Floppy disk
29 : COMPUTER: Macintosh
30 : OPERATING SYSTEM: Macintosh 7.0
31 : SOFTWARE: Microsoft word 5.1

```

	Query Match	3.4%	Score 33.6;	DB 2,	Length 44377;
	Best Local Similarity	57.7%;	Pred.No.2.1;	Mismatches 44;	Indels 0; Gaps 0;
	Matches 60;	Conservative			
QY	693 gatctgacacagatgagactcatatatgaataagaagagatgatcagtcgtttaga	752			
DB	34934 GAAGGACGCTTCCTTCTGTATGTGGCGCATCCGCCGTGCATGATGTTGGG	34993			
QY	753 tatgcagagatcttgatattgcatcttgtgttaattcccg	796			
DB	34994 TGTGCTGGAGGGTTTGGGGTGCCCGTGGCTGTGTTCAATG	35037			

```

1 Patent No. 6013455
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Bandman, Olga
6
7 APPLICANT: Tang, Y. Tom
8
9 APPLICANT: Hillman, Jennifer L.
10
11 APPLICANT: Yue, Henry
12
13 APPLICANT: Guegler, Karl J.
14
15 APPLICANT: Corley, Neil C.
16
17 APPLICANT: Gorgone, Gina
18
19 APPLICANT: Azimzal, Yalda
20
21 APPLICANT: Lu, Anna
22
23 TITLE OF INVENTION: Protein Kinase Homologs
24
25 FILE REFERENCE: PR-0614 US
26
27 CURRENT APPLICATION NUMBER: US/09/173,581A
28
29 CURRENT FILING DATE: 1998-10-15
30
31 NUMBER OF SEQ. ID NOS: 18
32
33 SOFTWARE: PERL Program
34

```

```

; SEQ ID NO 13
; LENGTH: 1866
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
;

```

Query Match	3.3%	Score 32.2	DB 3	Length 1866
Similarity	54.7%	Pred. No. 0.94		
Best Local				
Matches 64	Conservative 0	Mismatches 53	Indels 0	Gaps 0

```

1      RESULT 7
2      US-09-009-913-1
3      ; Sequence 1, Application US/09009913
4      ; Patent No. 6087485
5      ; GENERAL INFORMATION:
6      APPLICANT: AXYS Pharmaceuticals, Inc.
7      TITLE OF INVENTION: Asthma Related Genes
8      NUMBER OF SEQUENCES: 339
9      CORRESPONDENCE ADDRESS:
10     ADDRESSEE: Bozicevic & Reed, LLP
11     STREET: 285 Hamilton Ave, Suite 200
12     CITY: Palo Alto
13     STATE: CA
14     COUNTRY: USA
15     ZIP: 94301
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: Diskette
18     COMPUTER: IBM compatible
19     OPERATING SYSTEM: DOS
20     SOFTWARE: FASTSDS for Windows Version 2.0
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: US/09/009, 913
23     FILING DATE: 21-JAN-1998
24     CLASSIFICATION:
25     PRIOR APPLICATION DATA:

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEO-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEO ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 72928 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA

```

Query Match	3.28;	Score 31.4;	DB 3;	Length 72928;
Best Local Similarity	56.28;	Pred. No. 16;		
Matches 59; Conservative	0;	Mismatches 46;	Indels 0;	Gaps 0;

QY 318 aggtttagaagaatccagatgaatcaatcaatgaatgctgcgcgcgcgtatataaagcgagatgctgcgt 377
 Db 49094 AGGACACAGAGATCCAGCCACAGAGCAGTACGAGGATGATGATCAATAGGAATTTGGTAAAG 49153
 QY 378 cgatcaagctctgtagcgtatataagctgtgctgcgaattccgaatt 422
 Db 49154 GTAGATGAGCTATGTCGGCGGGAAGAGAGAGAGATTCAGATT 49198

```

RESULT      8
US-08-841-349-8/C
; Sequence 8, Application US/08841349B
; Patent No. 5955594
; GENERAL INFORMATION:
; APPLICANT: MISHRA, LODA
; TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT.
; FILE REFERENCE: XX/PO4470US0
; CURRENT APPLICATION NUMBER: US/08/841,349B
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (322)..(1509)
US-08-841-349-8

```

```

Query Match      3.2%; Score 31.2; DB 2; Length 1964;
Best Local Similarity 52.3%; Pred. No. 2.1;
Matches 69; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

```

QY 409 tcgaattcgaattcgtatgtatctaaacctgaacatccgagtggaattggt 468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 207 TAGCGCTTCATATAATATAATACATTCATTAATAATAATCAATAGTTCTG 148
QY 469 ggtaaagcatggaatggtctgtgagtcgtgaacacagaccggaaggtaccggat 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 147 TGGAAAAGGTGGATCTTTTTCGTGATGATACATACATGCTGAAGCATATTAATCT 88
QY 529 ttcatcatcattc 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 87 TAAATCATCTTCT 76

```

```

RESULT      9
US-09-033-055A-4
; Sequence 4, Application US/09033055A
; Patent No. 6069241
; GENERAL INFORMATION:
; APPLICANT: OHKAWA, HIDEO
; APPLICANT: IMAISHI, HIROMASA
; TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033,055A
; FILING DATE: MARCH 2, 1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PAUL E. WHITE, JR.
; REGISTRATION NUMBER: 32,011
; REFERENCE/DOCKET NUMBER: 9437/251563
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3651
; TELEFAX: (202) 822-0944

```

```

; TELLEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-033-055A-4

```

```

Query Match      3.1%; Score 30.2; DB 3; Length 1041;
Best Local Similarity 46.8%; Pred. No. 3.1;
Matches 95; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

```

QY 649 gttacgggtttcagaatgcagatggaatcaagatataaagatgctggaacagtg 708
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 266 GATAGCGTGCAGAAATATGAGATTCACAGCATGATTCATTCAGATGATGAGAAAGTG 325
QY 709 agaacatgatgatgaaggaagagatgatgctcagtgatgagtgatgagagatgctg 768
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 326 AAATTGGGAATTAAGAGAGAGTAAAGTGAATTGGAGAAATGCTGAGAGAAAGT 385
QY 769 ggaattcattgctgtgttaattcctgaagatctcaggtatcgaagccagatcga 828
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 386 GAATATATGGGAATATGCGAGAAAGCAAAATGATGATCGCGGAATATATAAGAAATGTA 445
QY 829 ggtttccgctgtgtctgataaa 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 446 AGGGCTGTGTATGTATGGAAGAA 468

```

```

RESULT      10
US-08-399-561-1
; Sequence 1, Application US/08399561
; Patent No. 5752903
; GENERAL INFORMATION:
; APPLICANT: Hirschberg, Joseph
; APPLICANT: Cunningham Jr., Francis X.
; APPLICANT: Gant, Elisabeth
; TITLE OF INVENTION: Lycopene Cyclase Gene
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5792903thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: MI
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,561
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4928 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS

```

LOCATION: 2029.3261
US-08-399-561-1

Query Match 3.0%; Score 29.6; DB 1; Length 4928;
Best Local Similarity 52.4%; Pred. No. 12;
Matches 89; Conservative 0; Mismatches 79; Indels 2; Gaps 1;

QY 783 tgggttaattcctgaagattcgaaggttttccgaacacgaacgaaggttcgcgtgt 842
DB 1663 TGCTGCTCAGCGTCGAACATTCGCCGCGTACCGAAGAGAGAAATCGCTGCATTCGCTGA 1722
QY 843 tctgaataagcctcctaagcttcgaggttcggttcgttcgagcggttcgagactcgt 902
DB 1723 CCGCGCAAGGCGCTCGAGCTGTCGAT--CTGTTTGACACCGGCGTCCATCCAGCGG 1780
QY 903 tgaagaagaatgatagaagcgttcgttcgaggaagaacccaagaac 952
DB 1781 TAAAGCCCTATTCTGCTGCTAGCGGTGAGTGTGTAAGAAAGTGAACAC 1830

RESULT 11
US-09-147-522-5/C
Sequence 5, Application US/09147522
Patent No. 6107069
GENERAL INFORMATION:
APPLICANT: MAGAGNIN, SIMONA
APPLICANT: BENATTI, LUCA
APPLICANT: CINI, MASSIMO
APPLICANT: SPECIALE, CARMELA
APPLICANT: COVINI, NEVIE
TITLE OF INVENTION: RECOMBINANT KYNURENINE-3-HYDROXYLASE ENZYME AND
FILE REFERENCE: 0769-0408-0PCT
CURRENT APPLICATION NUMBER: US/09/147,522
EARLIER FILING DATE: 1999-01-14
EARLIER APPLICATION NUMBER: PCT/EP/03589
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 5000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1507)
US-09-147-522-5

Query Match 3.0%; Score 29.6; DB 3; Length 5000;
Best Local Similarity 53.4%; Pred. No. 13;
Matches 62; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 601 gttctgtaaacactccgagatataacagcgttaagagagtgatgggttacgggttg 660
DB 4309 GTGCAGATTGAAGCTTGAGGAGTTACAAAGTATTGAATGGTCTACATATTCGGGTAA 4250
QY 661 ttgaatgcgcatgagatcagaagataaagatgattgtgaagaagtgagaaactga 716
DB 4249 TAAAGTTTGAATGATTAATAAAGCATTAAGCTCTGGGCGATTGAGTAGATCTGA 4194

RESULT 12
US-08-724-394A-20
Sequence 20, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston

APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-Oct-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fites, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.COMTIG"
US-08-724-394A-20

Query Match 3.0%; Score 29.6; DB 2; Length 246240;
Best Local Similarity 52.4%; Pred. No. 1.2e+02;
Matches 65; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 40 cttctctcatcatctctctcaaaaagactcaatatcttcccaagattcgtcaat 99
DB 152799 CATCTTCAACTAACTTCTCTTACTCAATTCCAATATTTTCATCAGGATTCATACAT 152858
QY 100 aacctagcagcagagagcagatcagatccgltcttaatttaacgcgaacgcgaactc 159
DB 152859 AATCAGCAATGGTGATTTTAAATTAAGAGAGATTGAGTAAATAGCAGGCTGACCA 152918
QY 160 gccg 163
DB 152919 TCCG 152922

RESULT 13
US-08-724-394A-21
Sequence 21, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereo

RESULT 14
 US-08-724-394A-22
 : Sequence 22, Application US/08724394A
 : Patent No. 5872237
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Feder, John N.
 : APPLICANT: Krommal, Gregory S.
 : APPLICANT: Laufer, Peter M.
 : APPLICANT: Ruddy, David A.
 : APPLICANT: Thomas, Winston
 : APPLICANT: Tsuchihashi, Zenta
 : APPLICANT: Wolff, Roger K.
 : TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
 : TITLE OF INVENTION: Sequences and Antibodies Thereeto
 : NUMBER OF SPOUNCES: 31
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
 : STREET: Two Embarcadero Center, 8th Floor

	Query Match	3.0%	Score 29.6:	DB 2:	length 246240;
	Best Local Similarity	52.4%;	Pred No.1.2e+02;		
	Matches	65;	Conservative	0;	Mismatches 59; Indels 0; Gaps 0.
QY	40	cttctccttcacatcatctctctcaacaagaagctcaatactctccccaagaattcgtaaat	99		
Db	152799	CATCTTCAACTAAACTTTCTCTTAAGTCACAAATTTCCTTTTTCATCAGAGATTCTAACAT	152858		
QY	100	aaccttagcgaacgagagtgccaatacagatcgcgtttcttcaattaaatcgcaaacgcggaaac	159		
Db	152859	AATCGACAATAAGTGATTTAATAAAAAGAGAGATTGTAAGTGAATAAATAGCAGCCTGAAACA	152918		
QY	160	gcccg	163		
Db	152919	TCCGG	152922		

US-RESULT 15
US-08-966-388-5
Sequence 5, Application US/08966388
Patent No. 5965412
GENERAL INFORMATION:
APPLICANT: TOMOYUKI NISHIMOTO
APPLICANT: MICHIO KUBOTA
APPLICANT: HIROTO CHAEN
APPLICANT: TOSHIO MIYAKE
TITLE OF INVENTION: KOJIBIOSE PHOSPHORYLASE, ITS PREPARATION AND USES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2001, 08:31:03 ; Search time 1361.86 Seconds

(without alignments)
6082.013 Million cell updates/sec

Title: US-09-553-431-3

Perfect score: 1182

Sequence: 1 aaagctgataatcgcaactcc.....tcaaaaaaaaaaaaaaaaaa 1182

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estbda:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estcov1:*
68: em_estcov2:*
69: em_estp11:*
70: em_estp12:*
71: em_estp13:*
72: em_estp14:*
73: em_estp15:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*
110: gb_est64:*
111: gb_est65:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estp16:*
 118: em_estp17:*
 119: em_estp18:*
 120: em_estp19:*
 121: em_estp20:*
 122: em_estp21:*
 123: em_estp22:*
 124: em_estp23:*
 125: em_estp24:*
 126: em_estp25:*
 127: em_estp26:*
 128: em_estp27:*
 129: em_estp28:*
 130: em_estp29:*
 131: em_estp30:*
 132: em_estp31:*
 133: em_estp32:*
 134: em_estp33:*
 135: em_estp34:*
 136: em_estp35:*
 137: em_estp36:*
 138: em_estp37:*
 139: em_estp38:*
 140: em_estp39:*
 141: em_estp40:*
 142: em_estp41:*
 143: em_estp42:*
 144: em_estp43:*
 145: em_estp44:*
 146: em_estp45:*
 147: em_estp46:*
 148: em_estp47:*
 149: em_estp48:*
 150: em_estp49:*
 151: em_estp50:*
 152: em_estp51:*
 153: em_estp52:*
 154: em_estp53:*
 155: em_estp54:*
 156: em_estp55:*
 157: em_estp56:*
 158: em_estp57:*
 159: em_estp58:*
 160: em_estp59:*
 161: em_estp60:*
 162: em_estp61:*
 163: em_estp62:*
 164: em_estp63:*
 165: em_estp64:*
 166: em_estp65:*
 167: em_estp66:*
 168: em_estp67:*
 169: em_estp68:*
 170: em_estp69:*
 171: em_estp70:*
 172: em_estp71:*
 173: em_estp72:*
 174: em_estp73:*
 175: em_estp74:*
 176: em_estp75:*
 177: em_estp76:*
 178: em_estp77:*
 179: em_estp78:*
 180: em_estp79:*
 181: em_estp80:*
 182: em_estp81:*
 183: em_estp82:*
 184: em_estp83:*
 185: em_estp84:*
 186: em_estp85:*
 187: em_estp86:*
 188: em_estp87:*
 189: em_estp88:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	295	25.0	710	93	AM668240	AM668240 GA_Fa001
2	253.6	21.5	873	94	AM729746	AM729746 GA_Fa002
3	190	16.1	557	107	BE35432	BE35432 DGI_40_A0
4	187.6	15.9	456	28	AL370067	AL370067 MtbA35D01
5	131	11.1	588	107	BE359345	BE359345 DGI_40_A0
6	129.8	11.0	773	108	BE418872	BE418872 SCL083.D0
7	126.2	10.7	278	40	AM160136	AM160136 pbl112 so
8	122.6	10.4	276	174	B23006	B23006 P16H1TF IGF
9	122	10.3	362	104	BE124676	BE124676 EST9371I
10	97.2	8.2	457	140	C99848	C99848 C99848 YAC
11	73.4	6.2	386	92	AM574469	AM574469 707092H09
12	72.8	6.2	1101	190	CNS0039C	AL063921 Drosophila
13	61.6	5.2	1146	191	CNS021G2	AL176843 Tetradon
14	58	4.9	1101	190	CNS00LT2	AL078714 Drosophila
15	57.8	4.9	1101	190	CNS000D1	AL065414 Drosophila
16	57.2	4.8	1101	190	CNS00LOO	AL068607 Drosophila
17	57	4.8	1101	190	CNS0100X	AL098379 Drosophila
18	55.2	4.7	919	190	CNS005RL	AL061409 Drosophila
19	54.8	4.6	767	190	CNS00AOX	AL055924 Drosophila
20	53	4.5	1031	190	CNS00CF2	AL059199 Drosophila
21	52	4.4	895	190	CNS0071A	AL066286 Drosophila
22	51.2	4.3	622	29	AU060989	AU060989 AU060989
23	51.2	4.3	1100	190	CNS0161V	AL106189 Drosophila
24	51	4.3	878	190	CNS0187R	AL108993 Drosophila
25	51	4.3	1080	190	CNS00EPP	AL069494 Drosophila
26	51	4.3	1101	190	CNS00396	AL063911 Drosophila
27	50.4	4.3	402	36	AV408561	AV408561 AV408561
28	50.4	4.3	420	36	AV411765	AV411765 AV411765
29	50.4	4.3	422	36	AV411541	AV411541 AV411541
30	50.4	4.3	859	190	CNS0044Y	AL055406 Drosophila
31	50.2	4.2	1101	190	CNS00399	AL063914 Drosophila
32	50.2	4.2	1101	190	CNS0106X	AL098495 Drosophila
33	50.2	4.2	1101	190	CNS017RX	AL108171 Drosophila
34	50.2	4.2	1201	190	CNS016CN	AL106577 Drosophila
35	50	4.2	1101	190	CNS016XR	AL107337 Drosophila
36	50	4.2	1201	190	CNS015YI	AL106068 Drosophila
37	49.8	4.2	884	190	CNS006UO	AL065923 Drosophila
38	49.6	4.2	1204	190	CNS016E2	AL106628 Drosophila
39	49.4	4.2	187	39	AM085373	AM085373 wy62H02.x
40	49.4	4.2	1084	190	CNS00Z5S	AL097258 Drosophila
41	49.2	4.2	1003	190	CNS017UQ	AL108524 Drosophila
42	49.2	4.2	1201	190	CNS0107R	AL098625 Drosophila
43	49	4.1	939	190	CNS00CNG	AL059400 Drosophila
44	48.8	4.1	928	190	CNS00DKY	AL071865 Drosophila
45	48.8	4.1	1101	190	CNS00LXB	AL078903 Drosophila

ALIGNMENTS

RESULT 1
 AM668240
 LOCUS
 DEFINITION GA_Fa0013E22 Gossypium arboreum 7-10 dpa fiber library Gossypium
 ACCESSION AM668240
 VERSION AM668240.1 GI:7502620
 KEYWORDS EST.
 SOURCE Gossypium arboreum.

Db	567	AAGTCTAATGATGGTGATAGAGGACACGGCGTGTAGAAATGTGATGGATTAAGCATAT	626
Oy	646	taaatgctcttgaacagagttagtaactgattgataagggtgaagatgatgctcgt	705
Db	627	AAAGATGATGGGAACACAGATGAGACGGCTCTGATTTAGAGGGGAAGATATGATCCCGT	686
Oy	706	ctctgag 713	
Db	687	TCTGATG 694	
RESULT	3		
BE359432			
LOCUS	BE359432		
DEFINITION	DG1_40_A03.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA		
ACCESSION	BE359432		
VERSION	BE359432		
KEYWORDS	EST.		
SOURCE	BE359432.1 GI:9300989		
ORGANISM	Sorghum bicolor		
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Sorghum.		
REFERENCE	1 (bases 1 to 557)		
AUTHORS	Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt ,L.H.		
TITLE	An EST database from Sorghum: dark-grown seedlings		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Cordonnier-Pratt MM Department of Botany The University of Georgia Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 542 1805 Email: mmp Pratt@uga.edu		
	Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.		
	Seq primer: Polymix		
	High quality sequence start: 6		
	High quality sequence stop: 531		
	POLYA=NO.		
FEATURES			
source	Location/Qualifiers		
	1..557		
	/organism="Sorghum bicolor"		
	/db_xref="taxon:4558"		
	/clone_lib="Dark Grown 1 (DG1)"		
	/note="Organ: 5-day-old dark-grown seedlings; Vector: Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."		
BASE COUNT	116 a 127 c 174 g 140 t		
ORIGIN			
Query Match	16.1%; Score 190; DB 107; Length 557;		
Best Local Similarity	69.2%; Pred. No. 1.2e-37;		
Matches	288; Conservative 0; Mismatches 125; Indels 3; Gaps 2		
Oy	482	ccgagattatatactataagattgcttcctgcagagatgatgacccggatccgagtcacacccgaatt 541	
Db	1	CCGGCTTATCTCTATCAGCACTGCCCGCAGAGGTGATGCCGGGTTTGCACTGCCATT 60	
Oy	542	acaaccgctcaagaagccgatattagttataaacaccgatattctgcatgagagatgca 601	
Db	61	GCACCTGCAGAAAGACAGAGCTCGTTACCAACCCCTGACATTCAGGCTCTCCGTATGCT 120	
Oy	602	gataagattacaagccttgcttgaattgataatgataaggagataataaagattgtgaac 661	
Db	121	GACCGTGTGGCAGAGACTGTTGAGTGCAGATGCATCAAGATATCAAGATTATTGTCAAC 180	
Oy	662	agaagttagaactgattgataagggtgaaagatatatgctcagttcttgatgltcaagag 721	

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	181	CGAGTGGGCGACACCGTGGTGAAGGAGGAGGAGACATGATGTCAGACATTTGATGTCACAGAA	240									
QY	722	atgttgagatgttcacatg--ttgagtataccagagatlc-gaaagtatccggaagtacg	778									
Db	241	ATGCTTGCGGTTGGCCCTTCTTGCTGGTGGTGTCTTCAGAGATGCGGAGTAATCCGAGATACA	300									
QY	779	aataagaggttcctgcctgtgtltaaaagcctccgacatcttagaagagattgagcattgaq	838									
Db	301	AATAGGGGTGATGACATTGGTGTCTTCACAGACACCGGCCCAAGCCTGGCGGCTTGTCTGTGAC	360									
QY	839	caaggtccttgagagattggttgagcaagatagcatgaagctctgtgagtgtgagga	894									
Db	361	CAGGCTACTGGCGCATGTTGGTGAAGAAGATGTGTATGACAGCATGATCGTTGAGGA	416									
RESULT	4											
LOCUS	AL370067	456 bp	mRNA	EST	03-AUG-2000							
DEFINITION	MLBA35D01.F1	MLBA	Medicago	truncatula	cdna	clone	MLBA35D01 T5,	mRNA				
ACCESSION	AL370067											
VERSION	AL370067.1	GI:9669820										
KEYWORDS	EST											
SOURCE	barrel medic.											
ORGANISM	Medicago truncatula											
REFERENCE	Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;											
AUTHORS	Magnoliophyta; eudicotyledons: core eudicots: Rosidae: eurosids 1;											
	Fabales; Fabaceae; Papilionoideae; Medicago.											
	1 (bases 1 to 456)											
	Journet E.P., Crespeau, H., van Tuinen, D., Gouzy, J., Jaillon, O.,											
	Niebel, A., Carreau, V., Chataignier, O., Kahn, D., Glaninazzal-Pearson											
	, V., and Gamas, P.											
	Medicago truncatula ESTs from nitrogen-starved roots											
	Unpublished (2000)											
	Contact: Genoscope											
	Genoscope - Centre National de Sequencage											
	Bp 191 91006 EVRY cedex - France											
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr											
	Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de											
	Biologie Moleculaire des Relations Plantes-Microorganismes,											
	CNRS-INRA, Bp 27 31326 Castanet-Polosan Cedex, France (Email :											
	Mt-est@toulouse.inra.fr Website :											
	http://sequence.toulouse.inra.fr/Mtruncatula.html).											
FEATURES												
SOURCE	Location/Qualifiers											
	1..456											
	/organism="Medicago truncatula"											
	/cultivar="Jemalong"											
	/db_xref="taxon:3880"											
	/clone="MLBA35D01"											
	/clone_1fp="MLBA"											
	/tissue-type="root tips"											
	/dex_stage="harvested after 3 days of N-starvation"											
	/note="Vector: phuescript psk. Site.1: EcorI, Site.2:											

Db 216 GTGTGGCCGTGGACGCCGCGGCTCCGGAACCTGACCTGCTGCGGCTGGAG 275
 QY 290 aacgcgtcaattacaccgctgttgaagttctcaacggcgattgacagctgacaaagcc 349
 Db 276 AACGGGTGAACCTGACCGCGCGGACGTGCTCGCGGCGAGCTGCGCTGACCAAGCC 335
 QY 350 ctgattgtgataaagcgtgtgtaaatcgaattcgaattgttatttcaaaactagtica 409
 Db 336 CTGCTCGGACACCGCTCTCCGCGGCTCCACCTGCTCTGCTCTCAAGCCCGGCTCC 395
 QY 410 aaattgccttagatttgagggaagatttagtttgcttgatgacttaaaatagg 469
 Db 396 AACCTGCCCCCTGCTTCTGCTGCTCCACACCTGACCTGCTGCTGCTGCTGCTGCTG 455
 QY 470 caagaaggtgtccggaatttatactataagattgttcctggaagttatgacggagttc 529
 Db 456 GGGCGGACCGCGCGGCTTCATGCTATGCTACGCTGCTGAGGTGTTGATGCTGGGTTT 515
 QY 530 ataccgcaattacacgcggtacgaagcggtattagttacaacactgatatatgca 589
 Db 516 GTGACAGCCATGCGCCCTGCGAAG-AGCAGTGTGCTGACACCTGACATGAC-6GA 573
 QY 590 ttgagagtcagataagttacaggttgcttgatgattgagtaagtaagataataa 649
 Db 574 CTCCTGATGACAGACCGCTGCGGCGGCTGTGTGATGTGACGCAATCAAGACATCAAG 633
 QY 650 at 651
 Db 634 AT 635

RESULT 7

AM160136 278 bp mRNA EST 01-MAY-2000
 LOCUS pbl12 soybean, century CDNA library Glycine max CDNA similar to
 DEFINITION SEPTUM SITE-DETERMINING PROTEIN, mRNA sequence.
 ACCESSION AM160136
 VERSION AM160136.1 GI:7673729
 KEYWORDS EST.
 ORGANISM Glycine max
 SOURCE soybean.
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Fabales; Fabaceae; Papilionoideae; Glycine.
 1 (bases 1 to 278)
 Matthews,B.F., Devine,T.E., Weisemann,J., Beard,H.B., Lewers,K.S., MacDonald,M.H., Park,Y.B., Malti,R., Lin,J.J., Kuo,J., Pedroni,M.J., Cregan,P.B. and Saunders,J.A.
 A soybean genetic map incorporating CDNA and EST markers and anchored with SSR markers
 Unpublished (2000)
 JOURNAL Contact: Beard HB
 Soybean and Alfalfa Research Lab
 USDA ARS PSI
 10300 Baltimore Ave, Bldg 006 Rm115, Beltsville, MD 20705, USA
 Tel: 3015045376
 Fax: 3015045378
 Email: hbeard@ars.arsusda.gov
 Insert length: 500 Std Error: 0.00
 High quality sequence stop: 278
 POLYA-No.

REFERENCE

AUTHORS

TITLE
 JOURNAL
 COMMENT

FEATURES

source
 location/Qualifiers
 1..278
 /organism="Glycine max"
 /cultivar="Century"
 /db_xref="taxon:3847"
 /clone_lib="soybean, century CDNA library"
 /note="Vector: pBluescript SK-"
 BASE COUNT 64 a 58 c 88 g 67 t 1 others
 ORIGIN

Query Match 10.7% Score 126.2 DB 40 Length 278;

Best Local Similarity 72.5% Pred. No. 1.4e-21;
 Matches 190; Conservative 0; Mismatches 69; Indels 3; Gaps 2;
 QY 662 agagttagaactgatttgataaggggtgaagatatgctgacgttcttgatgttcaagag 721
 Db 17 AGGGTTCGGACCGACATGATCAAGAGGAGACATGATGTCGTGTGACGTCTCAAGAG 76
 QY 722 atgttgagttgcatgtt--gattgatnccgagatcc-gaagtgtatcgagtagc 778
 Db 77 ATGTTAGGGTGGCTTGTGCTGCGGGGTATTCTGAGATAGTGAAGTTATTAAGACACC 136
 QY 779 aatagaagttccgctgtgttgaacaagcccgacttaagagattgacattgag 838
 Db 137 AATAGAGGTTTCCCTCTTGTGCTCAACACCTCCACTTTGGCCGATTTGGCTTGAA 196
 QY 839 caggtgtcttgagagatttgagcaagatagaagctgtgattgtgtgaggaagaa 898
 Db 197 CAAGCCGCTGAGAGCTCGAGCAAGATGATGACGCGCGGTGTGTGAAGAACAA 256
 QY 899 cctaaagaaggagatttctc 920
 Db 257 CCCAAGCTGGGTTTCTCT 278

RESULT 8

B23006 276 bp DNA GSS 10-OCT-1997
 LOCUS F16H1Tf tcf Arabidopsis thaliana genomic clone F16H1, DNA sequence.
 DEFINITION B23006
 ACCESSION B23006.1 GI:2507810
 VERSION B23006.1 GI:2507810
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 276)
 Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Coffey,E., Golden,K., Johnson,K., Adams,M.D. and Venter,J.C.
 A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing
 Unpublished (1997)
 Other-GSS: F16H1Tf
 Contact: Steve Rounsley
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: rounsley@tigr.org
 Seq primer: M13-21
 Class: BAC ends
 High quality sequence stop: 276.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source
 location/Qualifiers
 1..276
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="F16H1"
 /clone_lib="IGF"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
 Produced by Thomas Altmann"
 BASE COUNT 77 a 76 c 53 g 70 t
 ORIGIN

Query Match 10.4% Score 122.6 DB 174 Length 276;
 Best Local Similarity 75.6% Pred. No. 1.1e-20;
 Matches 152; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 754 agattggaagttgagtagacgaatagaggttccgctgtgtgtgaacaagctcc 813
 Db 754 agattggaagttgagtagacgaatagaggttccgctgtgtgtgaacaagctcc 813

Matches	241:	Conservative	0:	Mismatches	55:	Indels	66:	Gaps	2
QY	298	caatlacacgclglttgtaagtlctcaacgagcgttcagactcgaccgaagccctagtcg	357						
Db	1	CAATTACACCGTCTGTTGAGATTCTCAACGGGGAATGTCGACTCGAATCAAGCTTTGGTTAG	60						
QY	358	tgtataaacgcggtgcgaatttcgaattctctgtatcttcaaaccttaagtcanaattgac	417						
Db	61	AGATTAAACGTTGGTCTAATTTTGAATTCCTTGTAATTCCTAAACCTAGATTCGAATGATGTC	120						
QY	418	tttagagcttgggggaaagactttagtcttgccttgcctatcaaatgaatgaagcaaga	474						
Db	121	TATTGAGATTGGTGGAAAGCTTTGACTTGGTTGGTTGAAGCTTTAAAGCCCGGGGAGA	180						
QY	475	-----	474						
Db	181	TAAAACCGGTGAAATGGAACTCGAACCAAGCTTCTGCTCCGCTTCAGTTCCAGTTCG	240						
QY	475	--aggttgcgcggaattctatctatagatgtctgcgaagtatgtatgcgcggttcac	531						
Db	241	TTCTGCGCTGTCGCGAGCTTATTCCTGATCGATTGGCCCTGAGGAGATTGATCGCGGTTTAT	300						
QY	532	aaccgccttctacacgcgcttaacgaagccggtatagtttacaacctgatatctgcgat	591						
Db	301	AACGCGATTACACCGCGTAATGAAGCGGTTTGGATTACACACCGCGGATATTAACGCTT	360						
QY	592	ga 593							
Db	361	GA 362							
RESULT	10								
C99848		457 bp	mRNA	EST	08-OCT-1998				
LOCUS	C99848	YAC clone C1C8B11	region-specific cDNA	Arabidopsis thaliana					
DEFINITION	C99848	YAC clone C1C8B11	region-specific cDNA	Arabidopsis thaliana					
ACCESSION	C99848								
VERSION	C99848.1	GI:3719205							
KEYWORDS	EST								
SOURCE	thale cress.								
ORGANISM	Arabidopsis thaliana								
REFERENCE	Eukaryotic: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta								
AUTHORS	Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II								
TITLE	Brassicales: Brassicaceae: Arabidopsis.								
JOURNAL	1 (bases 1 to 457)								
COMMENT	Regional insertional mutagenesis of genes on Arabidopsis thaliana								
FEATURES	chromosome V using Ac/Ds transposon in combination with a cDNA								
source	scanning method								
BASE COUNT	Unpublished (1998)								
ORIGIN	Contact: Takuya Ito								
	Laboratory of Plant Molecular Biology								
	The Institute of Physical and Chemical Research (RIKEN)								
	3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan								
	Email: talto@ic.riken.go.jp.								
	Location/Qualifiers								
	1..457								
	/organism="Arabidopsis thaliana"								
	/db_xref="taxon:3702"								
	/clone_lib="YAC clone C1C8B11								
	region-specific cDNA"								
	112 a 125 c 96 g 122 t 2 others								
Query Match	8.2%	Score 97.2;	DB 140;	Length 457;					
Best Local Similarity	74.1%;	Pred. No. 3.2e-14;							
Matches 123;	Conservative 0;	Mismatches 43;	Indels 0;	Gaps 0;					
209	ctcgcgaggaacacccacacgagtcgtcgaacacgacgcgcgaagctgtctagtaaacctc	268							
Db	292	CTCGCTGCTTACCGGTTTCTCACTTCTGCGCATTTGACGCCGACCTGTGGTCTCCGTAACCTC	351						
269	gactctctctctcgtctcgaaaaacgcgcgcaattacacgcgtctgtgaagttctcaacgyc	328							

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR01J16"
/note="end : TET3"

BASE COUNT 469 a 6 c 69 g 151 t 406 others
ORIGIN

Query Match 4.9%; Score 58; DB 190; Length 1101;
Best Local Similarity 14.5%; Pred. No. 0.0003;
Matches 55; Conservative 173; Mismatches 152; Indels 0; Gaps 0;

Y 581 attactgcatgagagatgacagatgacagcttgccttgaatgagatgag 640
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 722 AKKATTTKKKKAAADKKRRKKGGKKGGKKGGKKGGKKGGKKGGKKGG 781
641 gatataaatgattgtgacagagatgacatgattgataagagtgatgatg 700
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 782 KAGDDADATKTTKKKKATTTTKKKKKKKKKKAAKADRTKTTKDAADAAA 841
Y 701 tcaattcttgatgacagagatggtgagatgcatgtgttgagataccgagatc 760
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 842 KTKDKKKKKKKTTTKKKKKKKKGGKKAAKKKKKKKKKKKKKKKKKKKK 901
Y 761 gaagatgctgagatgacagatgaggttgccttgccttgccttgccttgc 820
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 902 GKGKKDDAAAKKKKKKKKKGGKKGGKKGGKKGGKKGGKKKKKKKKKKAD 961
Y 821 gcaagatgcatgacagagctgcttgcagatgtgttgacagatgacagagct 880
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 962 AAKGTTKRAADAAADTDATKTTKKKKKKKKKKKKKKAAATKKKKKKKK 1021
Y 881 gtgagtgtgagagagacctaagagagagatcttgccttgccttgccttgc 940
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 1022 KKKTKKKKKKKKKADKKKKTKGKDKGKGGKKGGKKKKKKKKKKKKKK 1081
Y 941 tcaattcgtgacatgctg 960
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 1082 TADGKKKKKKAAADKKKK 1101

RESULT 15

CNS00001

LOCUS 1101 bp DNA GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR01J16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL065414

VERSION AL065414.1 GI:4938827

KEYWORDS GSS.

SOURCE

ORGANISM

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscophora; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2: cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR01J16"
/note="end : TET3"

BASE COUNT 280 a 104 c 123 g 211 t 383 others
ORIGIN

Query Match 4.9%; Score 57.8; DB 190; Length 1101;
Best Local Similarity 13.9%; Pred. No. 0.00034;
Matches 66; Conservative 221; Mismatches 188; Indels 0; Gaps 0;

Y 611 acagagcttgccttgaatgagatgagatgagatgagatgagatgagatg 670
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 627 WHAGTWTWKTKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 686
Y 671 actgattgataagagagagatgagatgagatgagatgagatgagatgagatg 730
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 687 AWKAGKAGAGKATKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAK 746
Y 731 ttgcatgttgccttgccttgccttgccttgccttgccttgccttgccttgc 790
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 747 KATAGAWMDARRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRGGRG 806
Y 791 ccgcttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 850
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 807 GGGGRRDARARARAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAKKAK 866
Y 851 agatgcttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 910
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 867 RWRDR 926
Y 911 ggaatttccttgccttgccttgccttgccttgccttgccttgccttgccttgc 970
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 927 GDMKRGTDWRTWKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 986
Y 971 gtttgccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 1030
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 987 RDTTDTAKRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1046
Y 1031 ttttgccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 1085
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 1047 KADAKKAKKADADADADADADADADADADADADADADADADADADAD 1101

Search completed: January 8, 2001, 08:31:08
Job time: 21508 sec

Qy 961 ttctctctcttgagcgc 978
 |||||
 Db 961 ttctctctcttgagcgc 978

RESULT 2

AB009056 65958 bp DNA PLN 09-AUG-2000
 LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MZF18.
 DEFINITION AB009056
 VERSION AB009056.2 GI:9758219
 KEYWORDS
 SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_1lb:Mitau1 P1

ORGANISM

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsids.

REFERENCE

1 (sites)
 Sato,S., Kaneko,T., Kotani,H., Nakamura,Y., Asanulzu,E., Miyajima,N.
 and Tanabe,S.
 and Tanabe,S.

TITLE

Structural analysis of Arabidopsis thaliana chromosome 5. IV.
 Sequence features of the regions of 1,456,315 bp covered by
 nineteen physically assigned P1 and TAC clones

JOURNAL

DNA Res. 5 (1), 41-54 (1998)

MEDLINE

98290546

AUTHORS

2 (bases 1 to 65958)
 Nakamura,Y.
 Direct Submission

TITLE

Submitted (27-NOV-1997) to the DDBJ/EMBL/GenBank databases.
 Yasukazu Nakamura, Kazusa DNA Research Institute, Department of
 Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935,
 Fax:81-438-52-3934)

COMMENT

On Aug 9, 2000 this sequence version replaced gi:2656032.

Address for correspondence: kazusaka@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kas/cgi-bin/agd-graph.cgi?c=MZF18
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Graal
 (Informatics Group, Oak Ridge National Laboratory,
 http://combio.ornl.gov/Graal-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hedsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremli1.zool.iastate.edu/cgi-bin/SP.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is M001 and the 3' clone is M188.

FEATURES

Location/Qualifiers

1..65958

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/chromosome="5"

/clone="MZF18"

/clone_1lb="Mitau1 P1"

join(355..1064,1116..5684)

/note="gene_id:MZF18.17"

/codon_start=1

/pseudo

/evidence-not_experimental

/product="non-LTR retroelement reverse transcriptase-like
 protein"

join(8240..8326,8416..8695,8789..9161,9246..9406,
 9571..9698,9776..10024,10109..10399)

CDS

/note="gene_id:MZF18.16"
 /codon_start=1
 /evidence-not_experimental
 /product="(+)-delta-cadinene synthase (d-cadinene
 synthase)"
 /protein_id="BAB08719.1"
 /db_xref="GI:9758220"
 /translation="MGSEVNRPLADPANIWEDPITSPKSLDGETFEKESHTLKEA
 VKRAPSRSKRNPIENIKFIDALGVSEKEDVLOLQSPCLDPPQVROEGC
 LYTGIIIFVYFROGFKLSADVEKFEKDEKGRFGHLITDQGLSLXEAQKSTHCE
 DIIDEALAFSRSHLEISRSRPHLARIRKALNHPYHKIGISLETQVISTYEEES
 CDPELSEFARTIDFNLQIILHREELACVIRWHEKPEFSKVYTRHRTTEALWELGY
 FEPQYSGARVITTMALITFALDQDEPIIPDSMKPIYVNVVEFQDIDELEWEGRS
 GGFHLKSLQKTANGYQOEAKMLKQDIIPFDEYKEMALISGVVALIATFVPMTD
 VAKLDAFEMLSHPKIRIVASEIISRFDDISSEFEKREHVAIGDIOCYOQFGVSK
 RAYVMGNIVSDAKMDLQNELMRPHVFPFLMLVNLRSRIDVFYRQDVTNPKLL
 KENHVSLLIETIP"

CDS

/evidence-not_experimental
 /product="acetyl-CoA:benzylalcohol acetyltransferase-like
 protein"
 /protein_id="BAB08720.1"
 /db_xref="GI:9758221"
 /translation="MKLELLSEVIRKSPASPHLQTLSTLDFDELPSTVYSAIFFYD
 HSNQEDIORLKSLSQTLSEFLYPLAGQIKDQVYHCNDGALFTEARKEIFLSDFLRN
 PSDADLQKFIYSPDHADPEFTWPLHVKVIFPKDGEVAVASV6KICDASLSTFVC
 SMTKASGVADTVDEFGADYFPADISIEFPLVHETSKTRKRFEGSLMIETLK
 NRASSGRVPQATRIESTTALLRCMTKAGSKGKVEFALITOMDPRVSSLLP
 HKAGNFPPLDLKESSEKWEIETVSKLOKTOELNELLRNQSEDDKSSVEKERI
 ASRALSLCEISPEMETRYANSKCRMSFEANFCGRVWVAPDSVDKTYQVLYNDSK
 SEGVARTLETETMAYREHSDLVIVTPSPSILI"
 complement(join(13438..13580,13695..14419,14516..14753,
 14956..15283,15400..15634,15951..16167,16240..16311,
 16400..16553))
 /note="gene_id:MZF18.14"
 /codon_start=1
 /evidence-not_experimental
 /product="PRO2 homolog"
 /protein_id="BAB08721.1"
 /db_xref="GI:9758222"
 /translation="MGNRSLSVLTMLVVLQWLLVWIMISTNLFKSKMTPLSKYLN
 TTYGPOSTNLVLTLPMPFIADVLSVYLHINKISQSLYELCEKMRKIRIGVIMV
 NMPGIYATELTSLFVALLANSLYLVLYSLYVHLHNDONAKIMQAFAPLRI
 GVCNIGCAFLPEPPTASTITPLVGLTSESSIKYHILGHSNCEFLVHYVFLIY
 AMINKMETPMPNPPYVNLGCTIMVYIGIMWVYSLSPFRKKEIFEPYHNLGLY
 IAFVYIHGDSMFCNLTNITLFLIDRLRLQSTKRSLVSAKILPQDNLELTSTK
 PGLHTPTSLIFLHPLSISKIOWHPTTITSSNLEKDLTVSVIRQGSNTOLLYHLS
 SGLDLEVTGPGVPGNSPDVSRHNSLLVSGSGITPFIIVIRELISQONKSTKLP
 DVLVCSFKRHADLAFDLIFPLDMSADISRLNRIEAYIREDKKEPTDNRHLO
 TKRFPQPLDPSIPVLGPNLELGVVLSFVFLIIGIVTYIYVVDHNGSI
 YNFSRGIDMDELGACIFISSVYELMRKONKGGDEKQVQVSVVCGPRKMHV
 MFGHERLESPYQSIQVATQVHSGSNPLKILLLEAGSDVGVVCGPRKMHV
 AKTSSGAKLHFAISFMV"
 complement(join(20836..20978,21097..21821,21910..22147,
 22283..22610,22727..22973,23265..23456,23532..23628,
 23708..23861))
 /note="gene_id:MZF18.13"
 /codon_start=1
 /evidence-not_experimental
 /product="PRO2 homolog"
 /protein_id="BAB08722.1"
 /db_xref="GI:9758223"
 /translation="MGNRSLSVLTMLVVLQWLLVWIMISTNLFKSKMTPLSKYLN
 TTYGPOSTNLVLTLPMPFIADVLSVYLHINKISQSLYELCEKMRKIRIGVIMV
 NMPGIYATELTSLFVALLANSLYLVLYSLYVHLHNDONAKIMQAFAPLRI
 GVRIGYGVCAFLPEPPTASTITPLVGLTSESSIKYHILGHSNCEFLVHYVFLV
 LIYAMVAKMETPMPNPPYVNLGCTIMVYIGIMWVYSLSPFRKKEIFEPYHNLGLY
 YGLTVFAIYVDSMFCNLTNITLFLIDRLRLQSTKRSLVSAKILPQDNLELTSTK
 PGLHTPTSLIFLHPLSISKIOWHPTTITSSNLEKDLTVSVIRQGSNTOLLYHLS
 TRLSSISLEVTGPGVPGNSPDVSRHNSLLVSGSGITPFIIVIRELISQONKSTKLP
 TKLPVNLVCAKRNHDLAFIDLIFPDSISVSDISKILNRIEAYITIREDKKEPTDDH

CDS
KLTGDKNRPODLSPISPYLGPNPNFLMGLVGLSSPYMFLILIGITVRYIYIPVDNR
TGSITNYFIRLYMFLMGLVGCVCIFISSISIFLMRKREKNEGDKKQVQSVERTPPIS
SPGSFHFHERELSESPYQSIQVQATSVHFGKPKMLKILFEAESESDGVWVCSPKRM
RHEVAKISSSLAKNLHREALSPFM*
Complement(Join(25651, 25974, 26145, 26386, 26718, 26876,
27106, 27216, 27389, 27652, 27695, 27963))

gene_id="MEF18.12"
/codon_start=1
/evidence_not_experimental
/protein_id="BAB06723.1"
/db_xref="GI:9758224"
translation="MALFLTVSPNPSPISPTTSLSHVBSRHHLLSSITIRNKSEPYAT
LPHALVSRRRRVKVSQSLRPLISPDHDHSSQMAALFAAAGFVSWSEKTKTIGSVGGL
ALTSLLILALASNNLRLLPFETPSYGFGEFFELPPTILLERADLRIRAIISGLSLA
FLIGSAIVAGTVAFAFMLVPMKSLGPDMMKTAALIMSGYIGSILNFAVIAISGLAISPA
VIAACVADNDYICALLHEFWLILALASKRIIPETASGSPADADTKDLKEDKRVVSTSI
ALASVEFLCKRAITLTTLTKTGQVMLPVATVITLVLASFPDENFNSKASAPSETSTLLIL
MNPETTLIAGGSVMNVNINFAISPIELISFAIOWHIAVTVLVGLFCDIMKLLALLASN
ANIGSEPTTACAMATKNGWTSILVPIGLISGVGVSLATFLIGCGCVFLKRL"
29181..31655

/codon_start=1
 /evidence=not_experimental
 /product="receptor-protein kinase-like protein"
 /protein_id="BAB08724.1"
 /db_xref="GI:9758225"
 /translation="MAPIINLTQTLIFECPCDLHLISFAATPTDNYILINSQNTSTSE
 FTTREFLSSESPGSSFLSDRSISITSPNNSDPSVPYLNARFVQSGFKPVTTHTE
 GHFRILHAPKAFKRNFLRSKFRVILNGSVISFSTYVAKEFLIKIDLPVLE
 SLPRKSGFVGNVAEVEGSAKODIMQKQRLVFNPAQSLNLSOVLFTVIRINVA
 GSKLTPTNDITLMTWVVDNTDLILAAARARMTHTSPNONGCATREIADNYIMTA
 QMDMDNDLOARENISNGFQVDERKYLIVLRHCDIVSSSLNNTLVNNTVETLRA
 KDVDLSTLEHTLASPLVIDPYAESDRSGMLRISGVPDLSNPNRYALLNGVETMR
 LSPVSEVSGAKRNVMLVQSGVDGFFLFLFSLVLCILRRKNNKTRSESGTMT
 LTRFNGSSRTKERTVSSGSHLTLLISFAKSTQNTNMFDSLVIGGVFVKPRGSL
 KNTREAVAKRSGSPSGRQGLPEFLSEITLITLKRHHLVNLVQSGCEMLLYEYKIM
 KQPLKSHLYGSTNPPLSKWOKRLEVICIGARGLVYHTSSQGIHRRDSTNILLDNK
 IYAKVAGLGRNSGPCIDETHVSTQSGFGLIDPEYRROOLDKDSVYSFGVALL
 VYCAPRAVDPLIVREOVNLAEMALQEMORSGMLIDQIQVDNIDELPCSLKFAETAEKE
 CCADYGRQPTIGDVIYMLNLEHVLQESGGLNIPEDYGVNIDTPCRAQGLNSGNSAE
 RYRGQGTSGSIISTQVFSQLMNAGR"
 33002 32982

/note="cds_id:30202.1"
 gene_id:MZF18.10"
 /codon_start=1
 /evidence=not_experimental
 /product="septum site-determining Mtd"
 /protein_id="BAB08725.1"
 /db_xref="gi:9758226"
 /translation="MASRLSTNHOSLLPESSLQKTLISSRPVNNPSRRSPRSRS
 LQPNRRPELAGETPRIVYVTSKGVGQVTTTANGSLAARGSVVADIDGILGRNGL
 DLILGNNRVNTICEVINGDCRDLQALVDRKMRNFELCIKSRKSRSLNPGFGKAL
 ENLVLDKTRPEGSPDFIILIDCPAGIDGFTIATIPANBAVLVTPDITALDADARV
 GLEEDKGIIDIMIVNRVETMDIKEDMMSVLDVDEMGLSLGVIPEDSEVIRSTNN
 GPPLYLNDPIAGLAIFECQAAARLVEDQSMKAVNVEEPRKRGFFSFFGG"
 complement(101(133889..34276,34361..34392,34402..35252
 35439..35984,36094..36264))
 /note="contains similarity to unknown protein
 gene_id:MZF18.9
 pirt||f002898"

	Query Match	Score	DB	Length
1	100.0%	100.0%	61	65958
2	100.0%	100.0%	61	65958
3	100.0%	100.0%	61	65958
4	100.0%	100.0%	61	65958
5	100.0%	100.0%	61	65958
6	100.0%	100.0%	61	65958
7	100.0%	100.0%	61	65958
8	100.0%	100.0%	61	65958
9	100.0%	100.0%	61	65958
10	100.0%	100.0%	61	65958
11	100.0%	100.0%	61	65958
12	100.0%	100.0%	61	65958
13	100.0%	100.0%	61	65958
14	100.0%	100.0%	61	65958
15	100.0%	100.0%	61	65958
16	100.0%	100.0%	61	65958
17	100.0%	100.0%	61	65958
18	100.0%	100.0%	61	65958
19	100.0%	100.0%	61	65958
20	100.0%	100.0%	61	65958
21	100.0%	100.0%	61	65958
22	100.0%	100.0%	61	65958
23	100.0%	100.0%	61	65958
24	100.0%	100.0%	61	65958
25	100.0%	100.0%	61	65958
26	100.0%	100.0%	61	65958
27	100.0%	100.0%	61	65958
28	100.0%	100.0%	61	65958
29	100.0%	100.0%	61	65958
30	100.0%	100.0%	61	65958
31	100.0%	100.0%	61	65958
32	100.0%	100.0%	61	65958
33	100.0%	100.0%	61	65958
34	100.0%	100.0%	61	65958
35	100.0%	100.0%	61	65958
36	100.0%	100.0%	61	65958
37	100.0%	100.0%	61	65958
38	100.0%	100.0%	61	65958
39	100.0%	100.0%	61	65958
40	100.0%	100.0%	61	65958
41	100.0%	100.0%	61	65958
42	100.0%	100.0%	61	65958
43	100.0%	100.0%	61	65958
44	100.0%	100.0%	61	65958
45	100.0%	100.0%	61	65958
46	100.0%	100.0%	61	65958
47	100.0%	100.0%	61	65958
48	100.0%	100.0%	61	65958
49	100.0%	100.0%	61	65958
50	100.0%	100.0%	61	65958
51	100.0%	100.0%	61	65958
52	100.0%	100.0%	61	65958
53	100.0%	100.0%	61	65958
54	100.0%	100.0%	61	65958
55	100.0%	100.0%	61	65958
56	100.0%	100.0%	61	65958
57	100.0%	100.0%	61	65958
58	100.0%	100.0%	61	65958
59	100.0%	100.0%	61	65958
60	100.0%	100.0%	61	65958
61	100.0%	100.0%	61	65958
62	100.0%	100.0%	61	65958
63	100.0%	100.0%	61	65958
64	100.0%	100.0%	61	65958
65	100.0%	100.0%	61	65958
66	100.0%	100.0%	61	65958
67	100.0%	100.0%	61	65958
68	100.0%	100.0%	61	65958
69	100.0%	100.0%	61	65958
70	100.0%	100.0%	61	65958
71	100.0%	100.0%	61	65958
72	100.0%	100.0%	61	65958
73	100			

Dd	32062	TCACAAAGAGACTCTAATATCTTACACCAAGATTGCAATTAACCCCTACGACGCGAGTGTCA	32121
Qy	121	atagaatccgctttcttcaaatltaatcgaacaacccggaatccgcggagaagacgcggtatc	180
Dd	32122	ATAGCATCCGTTCTTTCATATTATATGCAAAACCGGAACCTGCGGAGAAACGCGGATATC	32181
Qy	181	gtcgttatacactcccggaanaagcggtgtttgaaagacgaacaacacccgcaaatctcgt	240
Dd	32182	GTCGTTATCACTCCGGAANAAGCGGCTGTGGAAAGACGACAAACCCGAATTTCCGT	32241
Qy	241	ctctctccgcctcgttlaaggtlcccaagtctgtcgcacattgaacgcgcgacctgtgtccgt	300
Dd	32242	CTCTCTCGCGCTGTTACGGTTTCTCAGTGTGCGCATTTGAGCGCGACCTTGTCCTCGT	32301
Qy	301	aacctcgatctctcctccttggttggttaagaatcgaatgaatcaatctgtgcgaagtata	360
Dd	32302	AACCTCGATCTCTCTCTAGGCTTGAAGATGCAATGCAATTAACACTTGCTGCGAGTTATA	32361
Qy	361	aacggagaatttgcgtctcgtatcaagaactcctgtaagctgaagcgtttgtcgaatttcgaa	420
Dd	32362	AACGGAGATTGCGTCTCGATCAACCTCGTAGCTGATTAACGTTTGCGCAATTTGCAA	32421
Qy	421	tgtcgtatgatatctcaaaactatagatcagaacttccgaatggaatttgggttgtaaaacttg	480
Dd	32422	TTCGATGATGATTTCTTAACCTAGATCGAATCTCCGATGGATTTGGTGGTAAACATTG	32481
Qy	481	gaatgagcttgbatgactgttgaanaactagaaccggaaggtlccacgattlcaatcatalc	540
Dd	32482	GAATGAGCTTGTGATGATCGCTTGAANAACCTAGACCGGAAGGTTTCCACGGATTTCATCATATC	32541
Qy	541	gattctcccggaagatcgaatccgcgattcatcaaccgcattaccctccgcgcgaatgaaga	600
Dd	32542	GATTCTCCCGAGGAATCGATCCGGATTCTAATACCGCCATTACTCCGCGAATGAAGA	32601
Qy	601	gttctgtatacaactcccggaataataacagcgttaaggaatgctgataaggtlcccggttg	660
Dd	32602	GTTCTGTATCAACATCCGGATTATACAGCGTTAAAGGATCTGATTAAGGTTACGGGTTTG	32661
Qy	661	tlagaatcgaaatcggaatccgagataataagaatgattgttgaacaaagttgaactgatatg	720
Dd	32662	TTAGAATGCGATGAGATCGAGATATTAAGATGATGTGAACAGATGAGAACTGATATG	32721
Qy	721	atlaaaggagagatatgatatgcaggtgaagtgtgcgaagagatgcttgaagattgcatatg	780
Dd	32722	ATTAAGAAGGAGAGATATGATGTCAGTGTGATGATGTCCAGAGATGTTGGATTGTCAATTG	32781
Qy	781	cttggatgtaatctcttgaagaatcttgaagttatctcgaagcagaatccgaaggtlccgact	840
Dd	32782	CTTGCGTATATCTCGAATAATTCCTGAGTATTTCGAAGCAGCAATGAGAGTTTCCGCTT	32841
Qy	841	gttcgaaataaagccctccctcgccttgcgggaatttggcgttgaagaaagcgccttggaaagtc	900
Dd	32842	GTTCTGAATAAAGCTTCTACGCTTGCGGATTTGGCTTTGAGCAGCGGCTTGGAACATC	32901
Qy	901	gttgcgaagaatagatgaagaagcgttcatgtgtggaagaaacctaagaacgttgcttc	960
Dd	32902	GTTGAGCAAGATAGATGAAGCGCTGTATTAGTGGAGAAAGAACTTAAGAAAGTGGCTTC	32961
Qy	961	tctctcttcttggcggc 978	
Dd	32962	TTCCTCTTCTTTGGCGGC 32979	

RESULT	3					
AF251019	AF251019	1164 bp	mRNA	PLN	06-SEP-2000	
LOCUS	Tegetes erecta mind mRNA, complete cds.					
DEFINITION	AF251019					
ACCESSION	AF251019.1	GI:9971821				
VERSION						
KEYWORDS						
SOURCE	African marigold.					
ORGANISM	Tegetes erecta					
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta					

